THE UNIVERSITY
OF MICHIGAN
JAN 24 1961
BUSINESS ADMINISTRATION
LIBRARY

Journal of Advertising Research

Vol. 1, No. 2 DECEMBER 1960

Measuring Newspaper Readership by Telephone: Two Comparisons With Face-To-Face Interviews
DON CAHALAN 1

A Factor Analysis of the Liquor Preferences of French Consumers
JEAN STOETZEL 7

The Qualitative Values of Advertising Media
W. M. WEILBACHER 12

Another Look at Two Instant Coffee Studies
CONRAD R. HILL 18

Letters
JACK B. HASKINS 22
LESTER R. FRANKEL 23

FEDERAL STATISTICS IN ADVERTISING 25

RESEARCH IN REVIEW 26

PUBLICATIONS RECEIVED 34

EDITORIAL 36

Published by the Advertising Research Foundation

The Journal of Advertising Research

is published by the Advertising Research Foundation and sent free to its members. Single copies are available to eligible non-members of ARF at $10, and to ineligible non-members at $2.50.

solicits original papers. Reports of findings are favored over theoretical discussion. Manuscripts should be submitted in triplicate, double-spaced, with references, tables and figures on separate pages. Authors receive 50 free offprints. Letters of comment and criticism are also invited.

is intended for practitioners and users of advertising research. Limits on readers' time and journal space require that papers be as short as clarity permits.

is an open forum. Publication in it implies no endorsement of the writer's purpose, methods or views by the Advertising Research Foundation, its board of directors, or any of its committees.

is edited by the ARF Technical Staff:

Charles K. Ramond ..... Technical Director
Ingrid C. Kildegaard ..... Research Statistician
Gwyn Collins ..... Research Associate
Thornton C. Lockwood ..... Research Associate
George M. Shirey ..... Research Associate
Mary E. McMahon ..... Librarian
Joey Harris ..... Secretary

© 1960 by the Advertising Research Foundation, Inc.
3 East 54th Street, New York 22, N. Y.

Measuring Newspaper Readership by Telephone: Two Comparisons with Face-To-Face Interviews[1]

DON CAHALAN
ARB Surveys, Inc.

Do telephone and face-to-face interviews obtain roughly the same answers? The author found they did when a particular newspaper readership question was asked of new-car owners in and around New York.

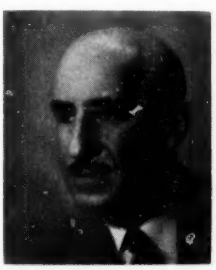
NEED FOR A TELEPHONE INTERVIEW technique to measure "yesterday" reading of weekday newspapers arose last year, when the New York *Herald Tribune* planned to survey the reading habits and characteristics of New York owners of late model automobiles. Since these owners were widely dispersed throughout five New York City boroughs and 23 counties within 50 miles, personal visits to a large, unclustered sample were considered too costly.

One established question about "yesterday" readership (Simmons, 1958) can be asked both in face-to-face and telephone interviews. The only differential effects one might anticipate would result from gestures or facial expressions of the interviewer or respondent. We doubted that such visual cues played a large role, but knew of no actual comparison of telephone and face-to-face interviews on "yesterday" newspaper reading. Two such comparisons were planned.

[1] The author wishes to thank the New York *Herald Tribune* for financial support of these studies, and Dr. Richard C. Sheldon, its research director, for assistance in all phases of the research.

METHOD

Both studies used essentially the same method. The interviewer asked the standard "yesterday" reading measurement questions over the telephone, and then arranged to see the respondent the same evening. In this follow-up face-to-face interview, the interviewer led the respondent through copies of all of the preceding weekday's issues of the seven major New York City daily newspapers, ostensibly



DON CAHALAN is president of ARB Surveys in New York. In 1937 he received a B.A. in journalism and in 1938 an M.A. in social psychology from the State University of Iowa. After four years in the Navy during World War II, he became associate professor of psychology and social science at the University of Denver. Since then he has held senior staff positions at the National Opinion Research Center, W. R. Simmons & Associates Research, the Bureau of Social Science Research, and, as a civilian, in the U. S. Army and Air Force. He has written a dozen articles on research methodology for various journals, including the *Public Opinion Quarterly*, *International Journal of Opinion and Attitude Research*, and *Journal of Medical Education*.

to find out which items looked "interesting." After the respondent had this opportunity to refresh his memory, the interviewer asked whether he was *sure* he did, or did not, look through that newspaper "yesterday."

If the respondent's face-to-face answer for any paper was the same as his response over the telephone, he was called consistent. The per cent of respondents who were consistent is reported for each paper, along with its face-to-face and telephone readership percentages.

Study I

This 169-interview study compared results from two subsamples of respondents: 71 interviewed first over the telephone and then face-to-face with scanning of newspapers the same evening, and 98 interviewed first face-to-face on the same "yesterday" reading questions as used on the telephone and again immediately thereafter, also face-to-face but with scanning of the newspapers. Both groups' second interview thus aided the respondent's recall of whether he had looked into or read any of the papers.

Both groups were also asked the same "warm-up" questions: which of the seven newspapers was regarded as "best for the news" and "for the best columnists and features," and which newspaper was looked at most frequently "for the ads." The names of the papers were read to the telephone respondents in rotated order, and each respondent was asked to note down the names of the newspapers in that order. In the personal visits, the interviewer presented the names of newspapers in a typed list.

The wording of the "yesterday" reading question, identical for the two groups, was:

"Now we'd like to know what papers you read or looked into—either at home or away from home.

"I will take them one at a time. Have you read or looked into a *weekday—Not* Sunday—issue of (NAME OF PAPER) during the past week?"

(IF "YES" OR "NOT SURE" FOR ANY PAPER) "It is very important to us to know *when* the different papers are read. When was the last time you read or looked through a *weekday* copy of (NAME OF PAPER)?"

(ASK IF "TODAY") "When was the last time before today that you read a *weekday* copy of (NAME OF PAPER)?"

This sequence of questions has been used (Simmons, 1958) to obtain "yesterday" readership as volunteered by respondent.

Telephone respondents were told:

"The editors and publishers of several of the New York newspapers are very much interested in knowing what types of newspaper materials are especially interesting. This will require me to show you copies of the New York daily papers. I'm located close to you; I'd like to come over now and ask you these last couple of questions. (I'm not selling newspapers or anything else, of course.)"

The wording of the aided recall, through-the-newspaper recheck of "yesterday" readership was the same for both groups:

"The newspaper people are interested in knowing what kinds of stories people like to read. Here is a copy of (NAME OF PAPER, FOLLOWING SAME ORDER AS USED IN EARLIER PART OF INTERVIEW). As I leaf through this issue, will you please tell me which of the ten numbered pages look interesting to you, and which pages *don't* look particularly interesting.

"I'd like to know, from your looking through those papers now, whether you are sure you did, or did not, read any edition of any of them *yesterday*. If you are not sure, please look at the papers again. Please remember that some papers have early and late editions which may look somewhat different.

"How about the (NAME OF PAPER)? Are you sure you did, or did not, read or look through a copy of (NAME OF PAPER) *yesterday*?"

Standardization of Materials: Since the criterion of reliability was whether the respondent would recognize a copy of a newspaper as one he had read "yesterday," we tried to aid recognition as uniformly as possible from one interview to another. Each interviewer each day was furnished copies of the seven New York metropolitan dailies for the preceding day. Editions were those distributed on the date of issue; no "bulldog" editions (distributed the night before the date of issue) were used.

As in the 1958 Simmons study, each issue was stapled so that only ten numbered pages would be shown the respondent in asking which pages looked interesting. This was done to keep the interview short.

To minimize possible respondent confusion between editions, the ten numbered pages were selected from only those pages which appeared in *all* editions of that newspaper. Where possible, the most distinctive pages were selected: sports, book reviews, financial pages and the like. For any given day, all copies for each of the seven newspapers were prepared in uniform kits for each of the interviewers.

The limit of ten pages per newspaper probably did not hamper the respondent's ability to recognize it as one he read "yesterday." He was encouraged to look at the paper again if he liked, before saying whether he had or had not read it the day before.

Interviewer Selection and Supervision: Eleven interviewers completed from ten to 28 interviews each. All interviewers had worked on market research surveys previously, mostly with ARB Surveys.

All interviewers attended two 90-minute pre-survey briefings at the ARB Surveys office, where the head of the field department, the project director and the assistant project director discussed details

of the pilot study and answered questions. Interviewers also were provided with five pages of written instructions. Regarding the purpose of the study, they were told only that it sought to obtain information about car ownership and "to discover newspaper reading habits."

Daily contact was maintained with each interviewer by sending him specially prepared copies of the seven newspapers, and by telephone calls and notes. Interviewers were urged to call in whenever problems arose in the field. After the interviewing, supervisors telephoned a randomly selected ten per cent of the respondents to find out whether they had been interviewed in accordance with instructions.

Sampling and Interviewing Controls: We drew two randomly chosen groups of potential respondents from the same universe. To get fairly evenly matched groups of respondents under conditions permitting follow-up visits just after the telephone interview, interviewers drew clustered samples from street address ("criss-cross") telephone directories where possible (New York City boroughs and Westchester County). Interviewers were given specific street intersections and told to select every *n*th name on streets extending outward from that point. Where criss-cross directories were not readily available, as in certain areas of New Jersey and Nassau County, interviewers selected names from regular telephone books for specific towns or cities in those areas. They first selected a page at random, and then called every name in a certain position on every *n*th page.

There was one interviewer each in Manhattan, The Bronx, and Westchester and Nassau Counties; two in Brooklyn, and four in New Jersey. One interviewer worked in both Manhattan and New Jersey and another in both Manhattan and The Bronx. Each was instructed to alternate his interviews between the all-visit interview and the telephone-and-visit interview. If it developed that there were more interviews of one type the first evening, this was to be balanced off in the second evening's work.

Interviews were conducted when men were thought most likely to be home: 5:30-9:30 p.m. Tuesday through Friday, and Saturday from 1:30 p.m. on. Interviewers made up to two calls to contact households in the telephone sample, and one attempt to contact households in the all-visit sample. They kept a tally of successful and unsuccessful calls.

We anticipated that the telephone-and-visit interviews would be more difficult than the visit interviews. Greater efforts were exerted (especially in the last few days) on the telephone-and-visit interviews in order to get an approximately equal number of both types. Completed telephone-and-visit interviews totaled 71, and visit interviews 98. All were done August 11-15, 1959, covering a full week of weekday issues of the seven New York metropolitan dailies. Because the telephone-and-visit group took more time per completed interview, our goal of an equal number of interviews each day for each group was achieved only approximately. Roughly equal proportions of each type of interview were obtained, however, for the first three days and the last two days.

The completion rate for visit-only interviews was 36 per cent of the eligible households. (See Table 1.)

TABLE 1

COMPLETION RATES IN STUDY I

	Combined Telephone-and-Visit	*Entirely Personal Visit*
Interviewed	22%	36%
Refused (the combined telephone-and-visit group includes instances of failure to get a follow-up interview)	24	11
Qualified respondent not home; or nobody home	54	53
Total potentially eligible	100%	100%
Base	322	273
Ineligible households:		
No auto owned in household	35	24
No male head in household	22	12
Telephone disconnected	2	0
Total households listed as selected	381	309

Principal causes of attrition were: no automobile in family, person responsible for automobile not home at time of call, and refusals.

Attrition in the telephone-and-visit sample was greater, even with two calls: 22 per cent of eligible households yielded completed interviews. Many respondents would not consent to an immediate visit the same evening. Some thought the visit was a ruse to sell something, others appeared to feel they had answered enough questions already over the telephone, and some could or would not keep their appointments. Obviously, these problems would not apply in a survey conducted solely by telephone.

RESULTS

Respondents reported their "yesterday" readership first by unaided recall, and second by the through-the-paper aided recognition technique.

Table 2 shows the per cent of respondents who gave the same or different answers both times for each newspaper. The largest proportion of under- or over-staters for any paper was six per cent.

It is possible that greater attrition in the telephone-and-visit group may have excluded more persons who tend to be inconsistent in reporting newspaper readership. Although we felt response consistency would be about the same among those who refused to be interviewed as among others, we may speculate about alternative hypotheses.

Perhaps refusers in the telephone-and-visit sample would have been twice as inconsistent as those who *were* interviewed. If this were true of the entire 24 per cent who refused in the telephone-and-visit sample, this would lower the per cent who were consistent in the telephone-and-visit group from 95 to 92 per cent, or a drop of three percentage points.

Table 3 compares the first and second readership percentages for each paper. In no case did they differ by more than six percentage points. It should be noted in this comparison that one respondent's overstatement can be counterbalanced by another respondent's understatement, and vice versa.

One reservation regarding the research design should be examined: the possibility that within either the telephone-and-visit or all-visit group, individual respondents may have deliberately tried to be consistent in their first (unaided recall) and second (aided recall) responses on the readership items. Ideally, to prevent any conceivable interaction between individuals' first and second responses

TABLE 2

STUDY I: PER CENT OF RESPONDENTS WHO GAVE SAME AND DIFFERENT ANSWERS IN SUCCESSIVE INTERVIEWS

	Telephone Unaided Recall Followed by Face-To-Face Aided Recognition (Base = 71)			*Face-To-Face Unaided Recall Followed By Face-To-Face Aided Recognition (Base = 98)*		
Newspaper	*Same*	*No-Yes*	*Yes-No*	*Same*	*No-Yes*	*Yes-No*
A	94%	0%	6%	97%	2%	1%
B	96	1	3	99	1	0
C	97	0	3	91	6	3
D	93	4	3	92	4	4
E	94	3	3	98	1	1
F	96	1	3	94	3	3
G	98	1	1	95	3	2
Average	95	2	3	95	3	2

READ: of the 71 interviewed first by telephone with the unaided recall procedure, and then by a face-to-face through-the-newspaper recognition measure, 94 per cent gave the same answers on Newspaper A, none first denied readership and later claimed readership, and six per cent first claimed and later denied readership.

TABLE 3

STUDY I: PER CENT OF RESPONDENTS WHO CLAIMED TO HAVE READ EACH NEWSPAPER "YESTERDAY"

	Telephone-and-Visit Group (Base = 71)		*All-Visit (Base = 98)*	
Newspapers	*Telephone (Unaided recall)*	*Face-To-Face (Aided recognition)*	*Face-To-Face 1 (Unaided recall)*	*Face-To-Face 2 (Aided recognition)*
A	41%	42%	47%	47%
B	30	28	33	33
C	15	13	17	20
D	13	7	13	15
E	11	14	14	14
F	13	13	12	11
G	14	13	8	9

READ: of the 71 respondents in the telephone-and-visit group, 41 per cent claimed to have read Newspaper A "yesterday" when asked on the telephone. Of the same 71 interviewed face-to-face with the newspaper immediately thereafter, 42 per cent claimed "yesterday" readership. Since the sample was limited to male new-car owners in certain parts of New York, these readership estimates may not be projected to the New York metropolitan area.

would have required two completely independent aided recall and unaided recall samples. Our judgment was that the cost of several hundred additional interviews for this purpose was not warranted, especially because the unaided recall and aided recall questions had been carefully planned to minimize respondent awareness of any inconsistency between his aided and unaided recall responses on readership. The questions were quite different: the unaided recall technique required responses in terms of the *last* time the newspaper was read, without suggesting that "yesterday" readership was the real focus of interest. The subsequent aided recall technique first led the respondent through the paper on questions of reader "interest" and then, after raising the face-saving possibility that papers have early and late editions which may look somewhat different, he was asked if he were sure that he did or did not look through each of the papers "yesterday."

Evidence that few respondents could have made an attempt to be consistent in their responses is seen in the very close similarity in the results with the two methods. If there were some interaction between first and second responses, the two answers of individuals in the all-visit group might have been systematically more consistent than the responses of persons in the telephone-and-visit group, because the longer time span between the telephone-and-visit interview (one-half to four hours longer than for the single face-to-face interview) allowed the former group of respondents more time to forget their first (telephone) responses. However, it is seen in Table 2 that there is no systematic difference in the correspondence of the two responses for the two groups of respondents. Thus both the nature of the questions and the empirical evidence seem to indicate that the interaction between first and second responses was negligible.

Study II

A second pilot study was conducted in October 1959 to test the telephone method with a larger sample. The methods were essentially the same as in the first pilot survey, except that the sample of 200 was confined to a comparison of the telephone interview with "yesterday" readership responses obtained from the same individuals the same evening.

Twelve regular interviewers were assigned the names and addresses of 500 car registrants, who were a subsample from the New York area sample for the *Herald Tribune* study. This group of 500 consisted of equal numbers of owners of late models of three low-price cars distributed throughout Bergen, Essex, Somerset and Union Counties in New Jersey; and The Bronx, Brooklyn (Kings), Queens, Nassau and Westchester Counties in New York.

As before, after the respondent completed the telephone interview, the interviewer sought the respondent's consent to an immediate visit the same evening, ostensibly to find out which features and articles in yesterday's newspapers the respondent thought to be "interesting."

Many of the selected respondents lived so far (up to 12 miles) from interviewers' homes that too many follow-up interviews were being lost because of long travel time. Accordingly, we stopped drawing respondents from the regular sample after 47 interviews had been completed. The remaining 153 interviews were obtained from street address directories, as in the first pilot study, in The Bronx, Brooklyn, Queens and Westchester County. Potential respondents were screened to select owners of 1957 or later models of automobiles.

Completion rates: Table 4 presents the accounting of completed interviews in relation to assigned potential interviews for the 47 interviews which were done in conjunction with the main study combined with results for the 153 interviews in the supplementary sample.

TABLE 4

COMPLETION RATES IN STUDY II

Total households contacted	1,536
No answer after two telephone calls	238*
Total households answering telephone	1,298
Not eligible for interview (no 1957-1960 auto)	915
Eligible for interview	383
No answer after two telephone calls, assumed to own a 1957-1960 auto	70*
Total estimated to be eligible for interview	453
Completed both telephone and personal-visit interview	200
Refusal—telephone interview	72
Unavailable for telephone interview (e.g., out of town during period of survey)	29
No answer after two telephone calls—assumed to be eligible	70*
Refused personal-visit interview after completing telephone interview	42
Unavailable for personal-visit interview after completing telephone interview	40
Total estimated to be eligible for interview	453

*NOTE: out of the 238 we may expect that the number actually eligible in terms of owning a 1957-1960 car would be 383/1,298 or 29.5% of 238 or an estimate of 70.

Two hundred, or 44 per cent of those estimated to be eligible, were interviewed over the telephone and immediately thereafter during the interviewer's visit to the respondent's home. Again, many respondents balked at an immediate visit, suspecting

sales efforts or other impositions, and many were unable to be home for the visit that same evening.

This relatively low completion rate probably did not affect the accuracy of the comparison between telephone and face-to-face interviews. To do so, it would have had to eliminate persons who were more likely to minimize or exaggerate their readership claims in a telephone interview relative to their claims of readership in a face-to-face interview. The tendency to claim reading differently on the telephone and in personal interviews may or may not exist, but it is probably not correlated with causes of attrition.

Table 5 shows the proportions of individuals who gave the same answers in both interviews.

TABLE 5

STUDY II: PER CENT OF RESPONDENTS WHO GAVE SAME AND DIFFERENT ANSWERS IN SUCCESSIVE INTERVIEWS

	Telephone Unaided Recall Followed by Face-to-Face Aided Recognition (Base = 200)		
Newspaper	*Same*	*No-Yes*	*Yes-No*
A	99.5%	0.0%	0.5%
B	97.5	0.5	2.0
C	99.5	0.5	0.0
D	97.0	1.0	2.0
E	95.5	1.5	3.0
F	97.0	1.0	2.0
G	98.0	1.0	1.0
Average	98.0	0.6	1.4

READ: of 200 interviewed first by telephone with the aided recall procedure, and then by a face-to-face through-the-newspaper recognition measure, 99.5 per cent gave the same answers on Newspaper A, none first denied readership and later claimed readership, and 0.5 per cent first claimed and later denied readership.

TABLE 6

STUDY II: PER CENT OF RESPONDENTS WHO CLAIMED TO HAVE READ EACH NEWSPAPER "YESTERDAY"

	Telephone-and-Visit Group (Base = 200)	
Newspaper	*Telephone Unaided Recall*	*Face-to-Face Aided Recognition*
A	54.0%	55.0%
B	32.0	38.0
C	29.0	29.0
D	28.0	29.5
E	22.5	22.0
F	15.0	15.5
G	12.0	13.5

READ: of the 200 respondents, 54% claimed to have read Newspaper A "yesterday" when asked on the telephone. Of the same 200 when interviewed face-to-face immediately thereafter, 55 per cent claimed "yesterday" readership. Since the sample was limited to male new-car owners in certain parts of New York, these readership estimates may not be projected to the New York metropolitan area. To discourage guessing, code letters for some of the newspapers have been changed from those in Table 2.

Consistency for individual newspapers ranged from 95.5 per cent to 99.5 per cent, with an average of 98 per cent. There was a small, systematic tendency for telephone readership to be less than the follow-up face-to-face through-the-newspaper readership of the same respondents, but this difference was not statistically significant.

Table 6 compares the first and second readership percentages for each paper.

Clearly, average newspaper readership was virtually identical in the telephone interviews and personal-visit interviews.

CONCLUSIONS

Within the limitations of the sample design used, we conclude that:

1. "Yesterday" readership of daily newspapers is measurable over the telephone for considerably less time and money than personal-visit interviews.
2. Telephone measures of "yesterday" readership can be expected to be the same as through-the-paper readership measures obtained in visits in at least 95 per cent of individuals' responses on specific newspapers.
3. Telephone and face-to-face "yesterday" newspaper readership percentages can be expected to differ by no more than a few percentage points.
4. These findings apply only to the rather highly routinized and thus rather memorable readership of newspapers. They may or may not apply equally to the measurement of audiences of other media.

REFERENCES

GUTTMAN, LOUIS. Problems of Reliability. In *Measurement and Prediction*, Vol. 4 of *Studies in Social Psychology in World War II*. Princeton: Princeton University Press, 1950.

HOVLAND, CARL I., IRVING L. JANIS AND HAROLD H. KELLEY. Test "Reliability": Its Meaning and Determination. *Psychometrika*, Vol. 12, No. 1, March 1947, pp. 1-6.

HOVLAND, CARL I., ARTHUR A. LUMSDAINE AND FRED D. SHEFFIELD. Comparison of the Before-After and the After-Only Design of Experiments. In *Experiments in Mass Communication*, Vol. 3 of *Studies in Social Psychology in World War II*. Princeton: Princeton University Press, 1949.

SIMMONS, W. R., AND ASSOCIATES RESEARCH, INC. *Profile of the Millions (2nd Ed.). A Study of the New York Market and Readers of Seven Daily and Six Sunday Newspapers*. New York: News Syndicate Co., Inc., 1958.

A Factor Analysis of the Liquor Preferences of French Consumers

JEAN STOETZEL
Institut Français d'Opinion Publique

If complex behavior is caused by a few simple motives, they can be found in the patterns of that behavior. Professor Stoetzel illustrates this premise with a factor analysis which tentatively explains consumer preference for nine liquors in terms of their sweetness, price and regional popularity.

PRACTICAL REASONS ACCOUNT for the introduction of motivation research in the field of market study. When it becomes desirable to influence consumer behavior and attitudes, a simple assessment is no longer adequate and there is an obvious need for a deeper knowledge of their roots. The word "motivation" expresses this concept of psychological causality.

But the very concept of causality is by no means a simple one, as many Western thinkers have repeatedly observed. Aristotle felt that it was necessary to distinguish among four kinds of causes. In modern times, Malebranche and Hume have demonstrated that in the realm of psychology as well as in physics, we are utterly unable to comprehend the effectiveness of a given cause, and that—however vivid—our spontaneous intuitions of causality are hollow and meaningless.

It would be a mistake to dismiss these remarks by scornfully characterizing them as philosophical. It is most important to be fully aware of what we are searching for. In the perspective of behavioral psychology, when we face the matrix of simultaneous preferences expressed by a group of consumers or voters, what corresponds to the popular concept of causality is simply the unique configuration or pattern made up by the diverse phases of the stimulus situation, the characteristics of the subjects and their behavior. And it is quite significant, I believe, that in French or German, the word *motif* or *motiv* is used to convey that very idea of a pattern in the field of architecture or music. It appears that in psychology, too, the notion of motivation can be resolved conceptually into that of a pattern. Needless to say, however, such structures are never given in actual experience, but are only intellectual constructions.

This inevitably leads us to the concept of the model. A model is a system of relationships, mathematically expressed, apt to reach a high degree of complexity, and such that when experience has set the values of the parameters, the values of all the variates can be deduced in agreement with the nearly inexhaustible data of observation.

METHOD

Such models of preference for several items by a population are made available through factor analysis. At first it appears that the individual choices between the several items are correlated, so that given a choice by an individual, all his other choices

JEAN STOETZEL is professor of social psychology at the University of Paris (Sorbonne) and director of the Institut Français d'Opinion Publique (IFOP), which he founded in 1938. His publications include *Théorie des Opinions* (1943) and a study of the post-war opinions of Japanese youth, *Without the Chrysanthemum and Sword* (Columbia University Press, 1953). In addition to directing the Sociological Research Center in Paris, he edits *Sondages*, a French public opinion quarterly, and the *French Sociological Review*. M. Stoetzel has served as president of the World Association of Public Opinion Research (WAPOR), the French Psychological Association and the International Association of French-speaking Psychologists.

can be forecast within determined probability limits. Later the whole matrix of correlation coefficients can be rationally computed from an appropriate set of variates characteristic of the items preferred.

The intended goal has thus been reached; a complex system of relationships has been established through which individual preference judgments, not one at a time but simultaneously, can be deductively reproduced. The principles of this reproduction are the factors—abstract mathematical parameters computed without any personal interference by the analyst. The analysis may thus be called entirely objective in the sense that anyone making use of the same procedure will find the same results. This feature gives way, as will be made clear later, to the possibility of turning over the burden of computation to machines.

Human judgment, however, is needed at the next stage when assumptions are formulated concerning the significance of the factors. Still, those assumptions themselves can be operationally expressed and consequently can be experimentally tested. The conclusions of the study are quantitatively phrased. Not only are the preferred items ranked by their loadings in the factors, but also the contribution of each factor to the variance of the empirical data is quantitatively assessed.

The preceding statements obviously need concrete illustrations. For this purpose, we will make use of a body of data gathered in the course of a market survey made in February 1956 by the French Institute for Market Research (ETMAR).

Personal interviews were conducted February 9-21, 1956 with 2,014 adult men and women who constituted a representative cross-section of the French population selected by quota sampling methods. All information was obtained in the home, in 161 different localities, by interviewers habitually employed by ETMAR. A total of 1,442 completed interviews were obtained, for a completion rate of 70 per cent.

In the course of this survey, the following question was asked:

> Which of the following liquors do you personally like best? You are requested to classify them by ranking on top the one you like best, ending up with the one you like least.

(A card with the following items was shown the subjects.)

ARMAGNAC	KIRSCH	RUM
CALVADOS	MARC	WHISKEY
COGNAC	MIRABELLE	LIQUEURS

RESULTS

Product-moment correlation coefficients were computed between the nine simultaneous choices requested from each respondent (see Table 1). The ranks given to Armagnac and Cognac on the one hand, and to Kirsch and Mirabelle on the other, showed positive, rather high correlations (.37 and .25). Armagnac, Calvados, Cognac and Marc were negatively and rather strongly correlated (—.38 and —.39) with Liqueurs. Rankings of Kirsch, Mirabelle, Rum and Liqueurs were negatively correlated with the other rankings. Most correlation coefficients with Whiskey ranks were small, the most noticeable exception being the Whiskey-Liqueurs correlation, which was rather highly negative, —.24. To sum up, the whole matrix of the 36 correlation coefficients revealed a rather clear-cut pattern of preferences by French consumers.

We can now turn to the factor analysis of this matrix. The procedure used was Thurstone's centroid method. Three factors have been extracted.

The analysis was discontinued after the extraction of the third factor because the third residual matrix was such that two-thirds (22/36) of the figures were less than or equal to three standard errors of the corresponding correlation coefficients. The values for the correlations deduced from the three factors differed 21 times out of 36 from the actually observed correlation coefficients by a quantity less than or equal to three standard errors.

TABLE 1

CORRELATIONS BETWEEN RANKINGS OF THE NINE LIQUORS

	Armagnac	*Calvados*	*Cognac*	*Kirsch*	*Marc*	*Mirabelle*	*Rum*	*Whiskey*
Calvados	.21							
Cognac	.37	.09						
Kirsch	—.32	—.29	—.31					
Marc	.00	.12	—.04	—.16				
Mirabelle	—.31	—.30	—.30	.25	—.20			
Rum	—.26	—.14	—.11	—.13	—.03	—.24		
Whiskey	.09	.01	.12	—.14	—.08	—.16	—.20	
Liqueurs	—.38	—.39	—.39	.90	—.38	.18	.04	—.24

A NOTE ON FACTOR ANALYSIS

Factor analysis is a way of explaining the correlation between various measures by assuming a few underlying factors which influence them all.

For example, we may have a set of tests of abilities all based on very different tasks. Yet we might find that people who do well on one test generally do well on many of the others. This might lead us to suggest that, to some extent at least, the tests measure the same thing—the same factors. Factor analysis is the mathematical technique for finding out how many factors are needed to explain the agreement between tests, and how much each test measures each factor.

The first step is to determine the degree to which the tests agree with one another. For this we calculate the correlation coefficients between each pair of tests. These can be arranged in the form of a "correlation matrix" as Professor Stoetzel has done for preferences in Table 1 on the page opposite.

From this, by appropriate procedures, is derived a "factor matrix" which shows how each test is correlated with each underlying factor. Professor Stoetzel's factor matrix is given as Table 2 below.

It is easy to see how the correlation matrix of preferences is related to the factor matrix. The correlation between the preferences for two liquors is explained by the way each of them is correlated with underlying factors.

Take Cognac and Armagnac for instance. The correlation between them is made up like this:

Correlation between Cognac and Armagnac = (loading of Cognac on Factor I) (loading of Armagnac on Factor I) + (loading of Cognac on Factor II) (loading of Armagnac on Factor II) + (loading of Cognac on Factor III) (loading of Armagnac on Factor III) = (−.60) (−.52) + (−.17) (−.03) + (.42) (.14) = .3012 + .0051 + .0588 = .3651.

This compares with the reported correlation of .37.

Some of the other correlations are explained less well in terms of factor loadings. This may be due to further factors not extracted.

Table 2 gives the factor loadings for the different liquors.

TABLE 2

LOADINGS IN THE THREE FACTORS

Items	*Factor I*	*Factor II*	*Factor III*
Liqueurs	0.64	0.02	0.16
Kirsch	0.50	—0.06	—0.10
Mirabelle	0.46	—0.24	—0.19
Rum	0.17	0.74	0.97[1]
Marc	—0.29	0.66	—0.39
Whiskey	—0.29	—0.08	0.09
Calvados	—0.49	0.20	—0.04
Cognac	—0.52	—0.03	0.42
Armagnac	—0.60	—0.17	0.14

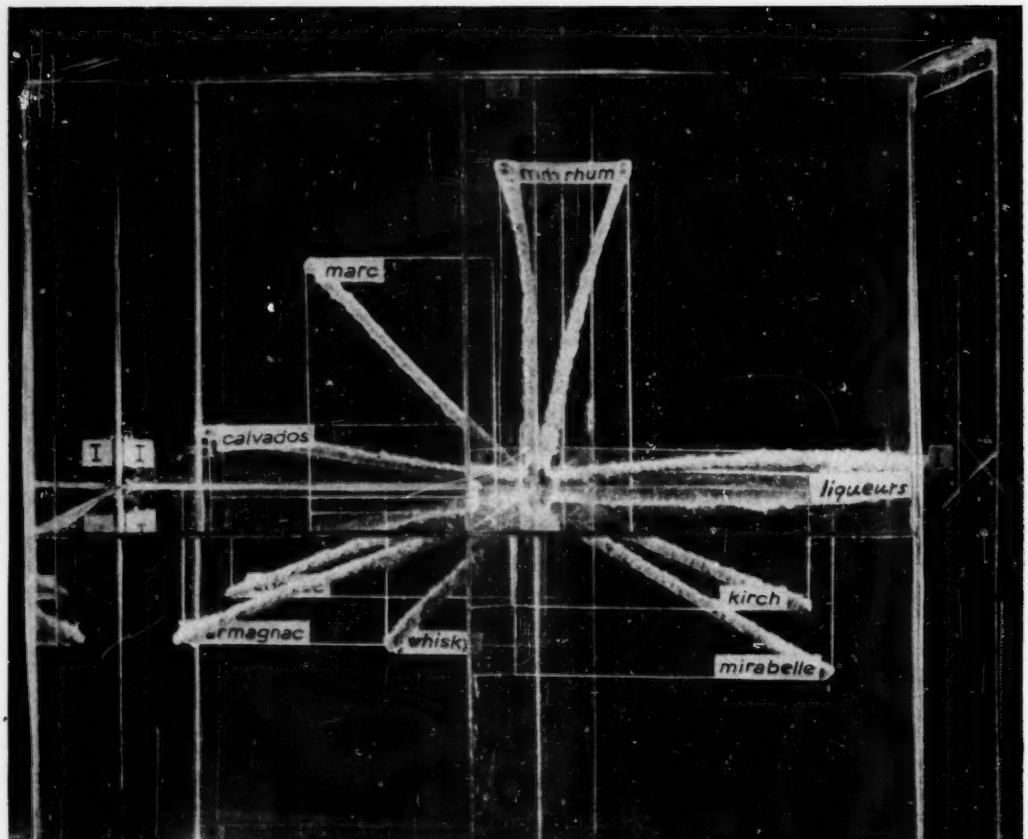
[1] This figure is certainly too high and should not exceed the maximum of .64. A reiteration of the factor analysis, bringing about progressively better approximations of the communalities, would very likely lessen it.

This table, and still more clearly a tri-dimensional geometrical representation, show again what was clear from the correlational matrix; the similarity in the rankings of Armagnac and Cognac, Kirsch and Mirabelle and also the main differences already observed. In addition, factor analysis puts those similarities and those differences in a complex order and reveals, in the abstract but strict sense of the word, *factors,* that is to say, the causes of these similarities and differences. Passing from the abstract to the concrete, from the mathematical to the psychological, is indeed delicate. We shall now proceed with this kind of interpretation by way of illustration.

DISCUSSION

The largest loadings in the first factor are, in order, those of Liqueurs, Kirsch, Mirabelle and the smallest those of Calvados, Cognac and Armagnac. We are inclined to surmise that the first factor discriminates between sweet and strong liquors. We conclude that this distinction between sweet and strong liquors is the first principle upon which consumer preferences are based. Let us emphasize this point: this principle is the first, not only in the sense that it has been analyzed first, but also in the sense that it is the most important. It is in effect this first factor which (except for Rum and Marc) contributes most to the variance in preferences.

Inexpensive



Strong *Sweet*

Expensive

A GRAPH OF THE FACTOR MATRIX

This is a photograph of a plastic model which graphs the factor matrix. Since there are three factors accounting for liquor preferences, the graph is three dimensional. The first axis runs across the page and represents, left to right, the dimension strong to sweet. Calvados, for example, is strong; Liqueurs are sweet. The second dimension is inexpensiveness and this runs up and down the page. Rum, for example, is inexpensive and exhibits plenty of this factor. The third dimension is less easily seen because it runs into the page. It represents local or national preference. The direction toward the reader shows local preference; national preference is indicated by positions away from the reader.

One interesting feature of this diagram is that it indicates the correlation between various preferences. Kirsch for instance enjoys about the same factor loadings as Mirabelle and the two liquors are, as we see from their neighboring positions in the model, positively correlated in preference.

TABLE 3

MEAN PREFERENCE RANKS BY REGION

Regions	*Armagnac*	*Calvados*	*Cognac*	*Kirsch*	*Marc*	*Mirabelle*	*Rum*	*Whiskey*	*Liqueurs*
The North	4.90	4.30	3.75	4.85	5.78	4.46	4.51	7.65	4.79
Normandy, Brittany	4.77	3.68	3.80	4.16	6.62	4.97	4.48	7.92	4.42
The Loire	4.42	4.62	4.16	4.23	6.75	4.15	4.82	7.94	4.41
Burgundy, Champagne	4.23	4.75	4.26	3.80	4.57	3.66	4.99	7.95	5.04
The East	5.21	5.87	3.92	3.51	5.92	3.65	4.29	7.81	4.65
West and Central	4.57	5.56	3.15	4.78	6.04	4.20	4.01	7.81	4.23
Aquitaine and Languedoc	3.62	5.53	3.32	4.86	6.36	5.17	4.06	7.65	4.35
Mediterranean and Alpine	4.36	5.49	3.99	4.85	5.45	4.95	4.63	7.90	4.24
Paris	3.92	4.89	3.64	4.16	6.25	4.14	5.25	7.01	4.74
Grand Mean	4.43	5.04	3.77	4.33	5.98	4.35	4.64	7.77	4.54
Variance	.2393	.2446	.1186	.2314	.3926	.2224	.1584	.0195	.0737

The second factor ranks the preference judgments in an entirely different dimension. This is quite natural since it is an essential property of the factors to be mathematically independent. The largest loadings characterize the preference judgment on Rum and Marc; the smallest on Armagnac and Mirabelle. A likely interpretation of this factor, which is given here only as a possibility, is that it is related to the price, low or high, of the different items. These differences in the prices of the several items, it may be observed in passing, need not be established objectively on the retail market. It is sufficient that they be so perceived psychologically by the respondents. This interpretation could be tested in a separate survey.

The third factor reaches its maximum in the case of Rum and Cognac, its minimum with Mirabelle and Marc. We have reasons to interpret it as related to the sociological variability of preference judgments within the public. A study of the mean rating given to each of the items throughout the different geographic regions of France (see Table 3) shows that the loadings in Factor III are negatively correlated with the variance of the mean ratings in the case of Marc, Mirabelle, Calvados, Kirsch and Armagnac, all of which enjoy a definite traditional preference in some region or another. In the case of the other liquors, for which regional differences are felt to a much lesser degree, it is likely that a more detailed study would show some other variable such as sex or socio-economic level to correlate with their loadings in Factor III.

CONCLUSION

This example may substantiate faith in the possible contribution of factor analysis to the problem of motivation. In the case studied, our interpretation of consumer behavior would be the following.

The major principle of liquor preference in France is the distinction between sweet and strong liquors. The second motivating element is price, which can be understood by remembering that liquor is both an expensive commodity and an item of conspicuous consumption. Except in the case of the two most popular and least expensive items (Rum and Marc), this second factor plays a much smaller role in producing preference judgments. The third factor is concerned with the sociological, and primarily the regional, variability of the judgments.

Consequently the recommendation of a systematic use of factor analysis in market studies seems justified. This procedure, as may have been noticed, leads to no change in the usual techniques of sampling and interviewing. The main reason why factor analysis has been used only on rare occasions in market research is probably that it requires rather laborious calculations. This difficulty is entirely overcome today thanks to electronic calculating machines. The preceding analysis was entrusted to l'Institut Européen de Calcul Scientifique of IBM-France. Once a program has been established, an IBM 704 electronic machine is fed with the ordinary punch cards of the market study at the rate of 150 per minute, during which time all significant data are read and enregistered and all arithmetic computations performed. Results are then printed at the speed of 100 lines a minute. Preliminary and final manipulations included, the working time was less than three-quarters of an hour.

We believe that factor analyses carried out under such conditions open up far-reaching and novel possibilities to market and opinion research.

The Qualitative Values of Advertising Media

W. M. WEILBACHER

Dancer-Fitzgerald-Sample, Inc.

A medium's qualitative value has been used to mean its audience composition, its editorial personality, or its contribution to advertising effectiveness. The author evaluates these definitions in terms of available research techniques, meaningful comparisons within and between media, and media selection.

THE ADVERTISING AND MARKETING practitioner is increasingly concerned with the evaluation of advertising media in terms of their qualitative or intangible values. This growing interest stems from two basic forces.

First, individual media and medium vehicles have had important successes in measuring quantitatively the characteristics possessed by both themselves and their direct competitors. Quantitative audience measurements are now well established for the major media, even though there are variations from medium to medium in the measurement bases used and in the number of measured vehicles. In addition, advertising message exposure data are increasingly subject to quantification, publicly for the printed media and, at least to some extent, in unpublished, private studies for the broadcast media. These message exposure studies are now severely limited in their coverage of individual medium vehicles, but future research studies will undoubtedly increase the knowledge base in this area.

W. M. WEILBACHER is vice president and director of research at Dancer-Fitzgerald-Sample and lecturer in advertising at Columbia University's Graduate School of Business. He received a B.S. in economics from Yale in 1949 and an M.S. in marketing research from Columbia in 1951. Following two years in the marketing research department of the National Biscuit Company, he joined Dancer-Fitzgerald-Sample in 1953 and served as manager of media research and research project director before assuming his present position. He is co-author, with Lyndon O. Brown and Richard S. Lessler, of *Advertising Media*, published by the Ronald Press in 1957. A member of ARF's Technical and Radio-TV Ratings Review Committees and of the 4A's Standing Committee on Research, Mr. Weilbacher also chairs the latter's Media Research Subcommittee.

Secondly, there is an increasing awareness of the need for media studies measuring the additional non-quantitative values accruing to an advertising message only because it is placed within a particular medium or medium vehicle. Such studies would measure the media values existing independently of the number of people to whom the message has been exposed.

These two forces are clearly implied by Alfred Politz in his introduction to the recent *Reader's Digest-Saturday Evening Post* exposure study:

> "Exposure is entirely the responsibility of the medium though the medium's function goes beyond that—into the mood it creates in its readers, viewers, or listeners, their confidence in the medium and other benefits the medium delivers for its advertising content. These other benefits are not measured in this study, which confines itself to exposure" (Politz, 1960, p. 8).

Increasing awareness of the precise size of media exposure potentials is leading to the desire for knowledge of what happens to an exposed person *as a result* of the placement of a particular advertising message in one medium rather than another, or in one medium vehicle rather than another. Such media effects are commonly referred to as qualitative or intangible media values. This paper discusses and analyzes three basic issues concerning the qualitative values of advertising media:

What is the qualitative value of an advertising medium or advertising medium vehicle, and how do currently available studies define and illustrate this value?

What is the significance of current research on qualitative media values from the standpoint of research methodology, inter- and intra-media comparisons, and effective media decisions?

What are the implications for future developments in media research?

QUALITATIVE VALUE DEFINED

There seems to be little published agreement among either advertising theorists or practitioners as to what, indeed, qualitative media value is. No single definition of qualitative value is common to most recent studies which either imply or state directly their purpose to be a measurement of qualitative value (Social Research, 1954; Social Research, 1959; Motivation Analysis, 1957; Coffin, 1959). This is in sharp contrast to the rather generally accepted definitions for audience membership underlying the basic quantitative media studies.

In informal or colloquial advertising industry usage, qualitative media value seems to connote some plus value of three specific kinds.

First, qualitative media value is frequently used to mean the qualitative characteristics of media audience members. These qualitative characteristics are merely descriptions of the audience attracted to the medium or vehicle by its editorial content. The value of these qualitative characteristics arises when they are demonstrated to be different from those possessed by the audiences of competing media or vehicles, or the general population.

There is no *direct* and *logical* connection between a description of the characteristics of an audience and the interaction of an exposed audience member with the medium or vehicle to which he is exposed. Nor do the characteristics of the medium's or vehicle's audience members directly and logically enhance advertisement effectiveness. The fact that an audience member is typically better educated or has more children, a higher income, or reads more hard cover books per year, etc., has not been shown to improve his receptivity to or appraisal of a medium or its advertising messages.

It is, however, still possible to build a logical (if hypothetical and indirect) case for the existence of a connection between audience member characteristics and the interaction between these exposed audience members and the medium or vehicle. Such an indirect case assumes that:

> If an audience of certain characteristics is developed by a medium or vehicle because of its editorial specialization, or whatever, then this specialized audience has a higher level of interest in specific subjects. It then follows that editorial content provides a more compatible environment for certain kinds of advertising messages and compels more attention to this environment from the audience members demonstrated to have unique characteristics. This interaction between the exposed audience member and the medium or vehicle should lead to greater receptivity to such advertising messages and to greater advertising effect. Such an argument can be carried to any kind of qualitative audience characteristic, but seems most powerful when based on audience qualities generated by a specific editorial specialization.

There is no proof that qualitative media values exist for the advertiser because of particular audience qualities. Certainly there seems every good reason to believe that audience quality must be of importance to a limited number of advertised brands (or advertised institutions) which, themselves, have specific and limited characteristics. The qualitative media value possibly caused by that audience quality would, however, vary from brand to brand (and institution).

A second meaning frequently given to qualitative media value involves the overt characteristics of the medium or vehicle itself and the way they are perceived by its audience. These qualitative values have to do with the role played by the medium or vehicle in the lives of the audience; with the audience member needs fulfilled by the medium or vehicle; and with the positive or negative attitudes toward it and its advertising created by the medium or vehicle in its audience.

In this sense, a qualitative media value is the value attributed by an audience member in terms of the real benefit or attitudinal benefit he derives from his contact with the medium or vehicle. Because media or vehicles tend to be very complex, they lend themselves to rather broad and involved description at this level. For example, in one newspaper study (Social Research, 1954, pp. 15-16) it was concluded, in part:

> "The newspaper is an integral part of the every-day life of a community. A citizenry identifies with the community via its newspaper. . . . Further, there is faith in the newspaper; this faith is a confidence in the integrity and believability of the content, and that what appears in its pages can pretty well be acceped as true. . . . It was seen that the people [audience members] who use newspaper advertising regard it as a *satisfying* kind of advertising . . . a kind of advertising *for them, and appropriate to their way of life. . . . Immediate,* and *close to the process of decision and purchase. . . . Sensible, easily understood* and *down-to-earth. . . .*"

Similarly, a radio study (Motivation Analysis, 1957, p. 25) concluded:

> "Listeners pay more attention to these CBS Radio stations than to the leading independents. Listeners distinguish between stations—regard these CBS Radio stations more favorably than the leading independents. Listeners believe these CBS Radio stations more than the leading independents—both their programs and their advertising. It seems it's a matter of authority."

These studies seek to establish and describe the kinds of interaction existing between the medium or vehicle and its audience.

This kind of qualitative media effect is defined in terms of words like "authority," "believability," "essential," "timeliness," "immediacy," "prestige," "mood," "provocative," etc. These translate into advertising effect if one can assume that the favorable

reception of the medium or vehicle generalizes to the advertising it carries.

That differentiated media personalities exist is, of course, one thing; that such personality factors relate to the advertising messages carried in the individual vehicle is quite another. The logic is clear and compelling, but there is no clear proof that it is correct. It remains to *establish* the precise relation of these media personality factors to advertising effect for individual brands and institutions.

Furthermore, the extent of advertising enhancement contributed by this kind of qualitative medium value must be different between advertising messages. The same editorial "authority," for example, must have an entirely different effect on cosmetic advertising as on that for earth moving equipment.

The third general usage of the phrase qualitative media value refers to the total positive or negative effect on audience member response to an advertisement because it has appeared in a particular vehicle. In this usage, qualitative media value is the total increment to advertising message effect contributed by the medium or medium vehicle. This qualitative medium value includes such individual contributors to overall qualitative value as the special characteristics of the audience attracted to the medium or vehicle and whatever personality characteristics are attributed to it by its audience members. Implicit in this third usage of the term qualitative media value is the notion that, beyond individual and definable sources of qualitative value such as demographic and attitudinal characteristics, the medium may also contribute completely undefined, even unanticipated, values to the advertising process. This point of view holds that it is wasteful to attempt to separate out and measure a complex pattern of myriad causes of medium or vehicle enhancement of advertising effect. Similarly, this point of view rejects the general measure of individual media characteristics preferring, rather, the specific measure of media effect on individual advertisements.

Thus, the objectives of a study conducted for *Fortune* are stated:

> "1. To determine whether the impact of an advertisement is modified by the business magazine in which it appears, and if so, in what ways it is modified;
>
> "2. To determine the degree of advertising impact achieved by three business magazines, *Fortune, Business Week,* and *U. S. News & World Report*" (Social Research, 1959, p. ii).

The logic in support of this point of view runs as follows:

> The concept of qualitative media value assumes that exposure to advertising occurs after exposure to a medium or vehicle has occurred. Once a medium or vehicle has achieved audience member exposure, any further value it contributes must be made in the context of the actual advertising message. This specific contribution almost certainly varies among different advertising messages (for example, as between cosmetics and earth movers). Therefore, in order to assess qualitative media value, its influence on the effect of a specific advertisement must be determined.

This usage of the term qualitative media value focuses on the relation of the advertisement to the audience member rather than on the relation of the medium or vehicle to the audience member. No indirect logic is required in this usage—it says, implicitly, "Qualitative media value exists; measure its effect on advertisements and make comparisons in terms of this measurement." The two previous usages of qualitative media value, on the other hand, require an intervening argument to link the measured quality of the medium with its advertising effect and fail to recognize the probability that qualitative media values will vary from advertisement to advertisement.

WHICH DEFINITION IS BEST?

These three usages of qualitative media value have served as bases for research published by advertising media. Interesting and valuable studies have been produced on the basis of each. At the same time they permit completely different conclusions concerning the qualitative potential of the same medium or vehicles.

This is, of course, to be expected. Each medium has its own strengths and weaknesses and it would be foolish for a medium not to explore and illustrate its own strengths. Similarly, it is in the best interest of the advertiser to understand the peculiar characteristics and strengths of each medium since his advertising goals are highly complex and variable, and since he must know which media are best adapted to each of his goals.

In this situation the advertiser must recognize that much qualitative media research is not pertinent to his own objectives or goals simply because it is research and describes particular peculiarities of particular media or vehicles. Rather, he must seek out those studies which bear directly on his own advertising efforts and it is, of course, in the best interest of the advertiser to insist that such research become increasingly less general.

In this context, it is pertinent to analyze each of the three points of view concerning qualitative media values presented above in terms of three criteria.

First, what potential and what limitations do each of these points of view have from the standpoint of research methodology? Do these concepts of qualitative media value lend or not lend themselves to strong research methodology?

Second, are the concepts of qualitative media value adapted to both intra- and inter-media comparison?

Third, what potential value does each of these studies have in the media selection process?

Audience Member Characteristics

1. *Research Limitations:* From the research standpoint, the description of audience member characteristics is, given an acceptable definition of audience exposure, straightforward enough. Using any of the current exposure definitions, it is simple enough to count those exposed and develop voluminous breakdowns of the various characteristics of audience members.

2. *Intra- and Inter-Media Comparison:* (a) When these measurement criteria are used as a basis for intra-medium comparison, vehicles within the same medium can be evaluated in a most straightforward manner. Many of the readily available media studies provide media data which are completely comparable in this sense.

It is, of course, the responsibility of the advertiser to examine such studies to assure himself that the exposure criterion used for measurement is germane to his own advertising goals. He should also scrutinize the study design to understand which characteristics, if any, of the media vehicles measured are highlighted or ignored, and to understand which of the measured media this tends to favor or short-change.

For example, the Nielsen Average Audience Measurement is now generally preferred to the Nielsen Total Audience Measurement in evaluating broadcast vehicle audience exposure to the average commercial contained within vehicles. The Total Audience Measurement tends to be rejected because, based as it is on a six-minute exposure, it produces progressively larger audience estimates as the length of the program increases regardless of the number of people exposed to the average minute.

(b) The usefulness of *inter*-media comparisons made on the basis of audience characteristics depends, on the other hand, upon how well exposures can be defined in equivalent terms for the various media. A detailed discussion of this subject is beyond the scope of this paper. Much work has been done with this controversial and difficult subject, and the most authoritative statement is undoubtedly that forthcoming from the Audience Concepts Committee of the Advertising Research Foundation.

Pioneering studies (e.g., MRCA, 1960; Politz, 1953) which have attempted to compare media on a common exposure basis have generally been subject to criticism on one or more of the three following grounds:

First, the same criterion for exposure is not used for each of the various media measured.

Second, the criterion for exposure is at such a general level that none but the broadest generalizations about an entire medium can be drawn. Although helpful, these generalizations frequently are not apt in specific media choice situations.

Third, the exposure measurement ignores all of the many other bases upon which choices between media are actually made (Brown, Lessler and Weilbacher, 1957, chaps. 13-15). Exposure is but one of the many characteristics and values offered by the various media; exposure data rarely (if ever) can be used alone to choose between media. To say that many of these other characteristics and values cannot or have not been quantified does not diminish their significance nor their importance in actual media choice situations.

3. *Value in Media Selection:* Such characteristic comparisons do not develop a direct connection between qualitative media value and incremental advertising effect. There is no clear quantitative measure of what or how much the medium or vehicle contributes to the advertising process independent of exposure. These measurements should prove helpful but not final guidance to individual advertisers in choosing between vehicles within media and give them limited insight into the nature of the differences existing between media.

Media Personality Studies

1. *Research Limitations:* The research difficulties are relatively more intense when it comes to the measurement of consumer attitudes and responses to media and media vehicles.

A review of typical studies of this type and independent pilot research in this area of consumer

attitudes and responses to media suggest the following, admittedly negative, methodological hypotheses:

First, consumers see little causal connection between media and the advertising messages which they carry. They think of a medium not as a carrier of advertising messages but as carriers of service, news and entertainment. Thus the viewpoint, reaction and total orientation of the consumer toward the medium is generally unrelated to advertising, to the interaction of advertising messages with the medium, or to the perspective of professional advertising people.

Assuming this hypothesis to be correct, *at most,* consumers can react to the editorial content of advertising media; most, if not all, attempts to develop consumer testimony as to the relation of media and copy, either generally or specifically, must be fruitless.

Second, consumers have almost no idea of their own relative involvement with a particular advertising medium or medium vehicles.

Assuming this hypothesis to be correct, variation in consumer involvement with media or the advertising messages they carry is not determinable from direct consumer testimony.

2. *Intra- and Inter-Media Comparisons:* Studies of this type tend to give the feel, the flavor, of a medium or vehicle. They do not, in their current form, give precise or quantified descriptions which permit objective comparisons.

Because these studies contribute most on the level of subjective appraisal through the delineation of individual medium or vehicle personalities, they are most helpful in comparing those media or vehicles which are least similar. Thus, this approach is likely most fruitful at the inter-media level, rather than in application to intra-media comparisons, since individual vehicles tend to be more, rather than less, alike.

3. *Value in Media Selection:* These studies of qualitative media values and their contribution to advertising effect are limited by their fuzziness, their lack of quantification and their essential dependence upon consumer introspection as a data source. Some steps are now being taken to overcome these failings through the use of techniques drawn particularly from projective psychology (Institute for Motivational Research Inc., 1960).

While these studies unfailingly give the flavor of the way media affect consumers, they generally do not, however, connect this flavor to the effect of specific advertising. This is left to the advertiser to appraise within the context of his own perspective, perceptiveness and problems.

Controlled Advertisement Effect Studies

1. *Research Limitations:* From a research standpoint this approach seems to have no essential limitation and great potential, particularly as classic experimental designs are adapted. The research possibilities seem to be limited only by the imaginations of the media researchers themselves. How much of a limitation this will be remains to be seen.

But which criterion of advertising effect is to be used? It seems likely that many of the traditional and experimental *copy research* approaches can be transferred to this work, since, after all, it largely involves the measurement of differential copy effect under varieties of controlled media conditions, copy held constant. What copy research measures might be adaptable is a subject beyond the scope of this paper, but certainly worth exploring.

2. *Intra- and Inter-Media Comparisons:* (a) All of the optimism in the preceding paragraphs applies directly to intra-media comparisons.

(b) It is doubtful, however, that this kind of comparison can be made *between* media, at least given the current state of the copy research art.

It seems unlikely that copy designed for and appearing in different media will be judged "equal" enough to permit controlled media testing. Unless prior agreement can be reached, for example, that television advertisement "A" is equivalent in communication content with printed advertisement "B," except for differences in effect contributed solely by the medium, it seems fruitless to attempt to measure the differences in effect contributed solely by the medium. It seems unlikely that practitioners or independent arbiters can reach agreement as to what constitutes "constant" copy as between media. Certainly, if a device is available for facilitating such agreement, it will be welcomed.

3. *Value in Media Selection:* Of the three approaches reviewed, the third type seems to offer the most promise of advancing our knowledge of how to make media decisions between vehicles of a given medium. It is most promising because it offers the opportunity of measuring qualitative media values at the level of the advertisement.

Inter-media comparisons of qualitative media value must, for the time being, apparently be made at the level of media effect on audience members. This kind of measurement can take the form of describing audience characteristics or of describing audience attitudes toward the medium, as discussed above.

IMPLICATIONS FOR MEDIA RESEARCH

This analysis suggests two basic conclusions for media research.

First, media research, as traditionally defined and practiced, may be ill-equipped to handle successfully measurement in the area of qualitative media effect. The orientation of large-scale consumer surveys aimed at identifying various audience characteristics—from audience membership to frequency of exposure to the average page—is important to the future development of media research. Equally important, however, is an orientation toward the effects of individual advertisements as measured by the qualitative techniques of modern-day behavioral science. Progress will depend upon an appreciation of experimental design and its potential contribution when combined with measurements of advertisement effect against population samples of medium vehicle audiences.

The typical divorce of media and copy or motivational research in advertising agencies, the frequent ignorance of copy or motivational research in the research departments of the leading media, and the late dawning of advertising research activity in many advertiser organizations have all combined against the development of completely rounded media researchers.

A basic viewpoint of this paper is that qualitative media value should be defined in relation to the overall effect of advertising and in interaction with specific advertising copy. Given this perspective, it is argued that research within this area must proceed not only from a theory of how media works, but also from a theory of how advertising copy works and how these two combine to work together. Media researchers have not traditionally been developed and trained to handle problems within these dimensions.

Second, expansion of media research beyond the important quantitative studies now available is inevitable. In the process, the melding of all current tangents of advertising research is also inevitable. The next developments in media research must come from increasingly rigorous measurement of qualitative media effect. If such studies, carefully conceived and executed with fully representative samples of medium or vehicle audiences, reveal important differences between media vehicles in their qualitative contribution to media effect, important gains in the planning of advertising efficiency will be made.

One can hope that media studies of the immediate future will be designed to demonstrate the presence or absence of qualitative media effect rather than merely reiterating the known competitive differences between media vehicles already absorbed into the media planning processes.

REFERENCES

AMERICAN RESEARCH BUREAU, INC. *The Local Television Audience* (a syndicated service published regularly). Beltsville, Maryland: American Research Bureau, Inc.

AMERICAN RESEARCH BUREAU, INC. *The United States Television Audience* (a syndicated service published regularly). Beltsville, Maryland: American Research Bureau, Inc.

AUDITS AND SURVEYS COMPANY, INC. *Exposure of Advertising—An Experimental Study*. New York: Time, Inc., 1959.

BROWN, L. O., R. S. LESSLER AND W. M. WEILBACHER. *Advertising Media*. New York: The Ronald Press Company, 1957.

COFFIN, THOMAS E. Total Effect Is Called the Key to Meaningful Media Comparisons. *Media/scope*. February 1959.

INSTITUTE FOR MOTIVATIONAL RESEARCH, INC. *The Advertising Effectiveness of Family Circle*. New York: Everywoman's Family Circle, 1960.

MARKET RESEARCH CORPORATION OF AMERICA. *The Profitable Difference: A Study of the Magazine Market*. New York: Magazine Publishers Association, 1960.

MOTIVATION ANALYSIS, INC. *Mike & Mike—are not alike—Different!* New York: CBS Radio, Inc., 1957.

NIELSEN, A. C., COMPANY. *Nielsen Radio Index* (a syndicated service published regularly). Chicago: A. C. Nielsen Company.

NIELSEN, A. C., COMPANY. *Nielsen Station Index* (a syndicated service published regularly). Chicago: A. C. Nielsen Company.

NIELSEN, A. C., COMPANY. *Nielsen Television Index* (a syndicated service published regularly). Chicago: A. C. Nielsen Company.

POLITZ, ALFRED, MEDIA STUDIES. *Ad Page Exposure in 4 Magazines*. New York: The Reader's Digest Association, Inc. and The Curtis Publishing Company, 1960.

POLITZ, ALFRED, MEDIA STUDIES. *A Study of Outside Transit Poster Exposure*. New York: National Association of Transportation Advertising, Inc., 1959.

POLITZ, ALFRED, RESEARCH, INC. *The Audiences of Nine Magazines*. New York: Look Magazine, 1958.

POLITZ, ALFRED, RESEARCH, INC. *Duplicate Magazine Coverage of Households and Consumer Expenditures, Volume Five, Life Study of Consumer Expenditures*. New York: Time, Inc., 1959.

POLITZ, ALFRED, RESEARCH, INC. *A Study of Four Media*. New York: Time, Inc., 1953.

POLITZ, ALFRED, RESEARCH, INC. *A 12-Months' Study of Better Homes & Gardens Readers*. Des Moines, Iowa: Meredith Publishing Company, 1956.

PULSE INCORPORATED, THE. *U. S. Pulse TV* (a syndicated servcated service published regularly). New York: The Pulse Incorporated.

PULSE INCORPORATED, THE. *U. S. Pulse TV* (a syndicated service published regularly). New York: The Pulse Incorporated.

SOCIAL RESEARCH, INC. *Advertising Impact and Business Magazines*. New York: Time, Inc., 1959.

SOCIAL RESEARCH, INC. *A Study of the Functions of Newspapers for Their Readers* (condensed). New York: American Newspaper Publishers Association, 1954.

Another Look at Two Instant Coffee Studies

CONRAD R. HILL
University of Nebraska

Professor Hill suggests that direct and indirect questions both have limitations in consumer research because answers to both are likely to be rationalizations.

JUDGING FROM THE MOMENTUM of published reports, the use of motivation research as an advertising tool has languished considerably since the faddish frenzy of the early fifties. Despite this gradual fade-out, one—Mason Haire's study of instant coffee—continues to stand out and earn kudos and footnotes from research writers (Selltiz *et al*, 1959). The purpose of this article is to assess this "classic" article (Westfall, Boyd and Campbell, 1957, p. 135) so far as practical advertising and marketing problems are concerned.

Haire's premise is that advertising research methods which rely exclusively upon direct questioning of consumers usually elicit mere rationalizations about product attributes. Therefore, in order to uncover the more important but latent attitudes, projective techniques are required.

Haire suspected that "dislike for the flavor of instant coffee" was just a ready rationalization. To test this notion he showed different grocery lists to two groups of 50 housewives each. Group One was asked to describe the housewife who made up the following shopping list:

> "Pound and half *(sic)* of hamburger
> Two loaves of Wonder bread
> Bunch of carrots
> One can Rumford's Baking Powder
> Nescafé instant coffee
> Two cans Del Monte peaches
> Five pounds potatoes" (Haire, 1950, p. 651)

Group Two was asked to describe a housewife who made up a list identical to that above except that "Maxwell House Coffee (Drip Grind)" replaced "Nescafé instant coffee."

Before recapitulating Haire's findings in Table 1, let's take a closer look at the suspected cultural stereotype that Haire discusses—dislike for the flavor of instant coffee. His respondents were asked, "Do you use instant coffee?" If the answer was "No," they were then asked, "What do you dislike about it?" Obviously this question introduces bias because it assumes that non-users are non-users due to dislike, which may or may not be true.

A better question would have been, "Have you used instant coffee within the past week?" or some other question that would have avoided the defensive sting of a direct accusation. What's more, Haire should have asked for likes as well as dislikes. Failure to look at the other side could load his data seriously.

CONRAD R. HILL is assistant professor of journalism at the University of Nebraska. After obtaining his B.A. in journalism from the University of Michigan in 1950, he worked as an advertising executive for Sparks-Withington, Grey Advertising and the Atomic Research Corporation. He received an M.A. in communications from Stanford in 1957 and has since taught at the Universities of Missouri and North Carolina. His first book, *The Aberrant Image of Advertising*, was published this year by the University of Nebraska Press.

As Haire's Table 1 shows, "flavor" as a determinant of respondents' underlying attitudes toward instant coffee failed to arise in the depth situation. Instead, attitudes were pitched in terms of value judgments about the shopper. This is not surprising because that is what he asked for. The instant coffee shopper is viewed as lazy and as failing to plan her household duties and purchases. On the other hand, the regular coffee woman shines with the positive, culturally approved, housewifely virtues. Haire's follow-up pantry check revealed that instant users tended to be more tolerant of the instant coffee shopper, whereas the non-users were more negativistic. From these findings, Haire concluded that these "underlying attitudes" influence the decision to buy or not to buy instant coffee, that the only way to unearth these important attitudes is by projective research, and that direct questions cannot bring forth the "real" information so necessary for effective advertising campaign appeals.

As this critique suggests, that is one whale of a supposition. Negative attitudes toward sinners do not prevent most of us from sinning; negative attitudes toward instant coffee may not prevent us from buying it.

Comments on Methodology

Haire's failure to describe his sample design makes it difficult to assess the validity of his computational methods. Overlooking this basic consideration, there are other methodological questions that may be legitimately raised. Even if we grant that attitudes toward instant coffee and instant coffee users are synonymous—which they are not—Table 1's presentation suggests other weaknesses.

Because the Table 1 presentation offers gross analysis, it doesn't really tell us very much. Needed are determinants of relationships between the attitudes of users and non-users by age, income, family size and so forth. This is impossible with a sample of 50, because there are not enough respondents for the cells demanded in cross-tabulations. In other words, it is statistically unwise to draw conclusions and inferences from two respondents per cell.

As for the pantry check, we learn there were 32 users and 18 non-users of instant coffee. Projecting that ratio to the universe as Haire did with other findings, are we to assume that instant coffee enjoyed 64 per cent market penetration in 1950?

Another methodological criticism involves the selection of classes in Table 1's stub. Do these meet the research criterion of being mutually exclusive? For instance, could it be that "good housewife, plans well, cares for her family" and "plans balanced meals" represent a single "planning" dimension, and not two as Haire suggests? In other words, do they overlap?

If so, then the findings may be heavily loaded with articulate respondents, that is, those who harped on the same theme with only slight variations. Additional discussion of this aspect follows.

TABLE 1

MENTIONS BY USERS AND NON-USERS OF INSTANT COFFEE
(AFTER HAIRE, 1950)

	By Women Who Had Instant Coffee in the House (N = 32)		*By Women Who Did Not Have Instant Coffee in the House (N = 18)*	
The woman who buys Nescafé is seen as:	*Number*	*Per Cent*	*Number*	*Per Cent*
Economical†	22	70	5	28
Not economical	0	0	2	11
Can not cook or does not like to†	5	16	10	55
Plans balanced meals*	9	29	2	11
Good housewife, plans well, cares about family†	9	29	0	0
Poor housewife, does not plan well, does not care about family*	5	16	7	39
Lazy*	6	19	7	39

*An asterisk indicates that differences this great would be observed only five times out of 100 in repeated samplings of a population whose true difference is zero.

†A dagger indicates that the chances are one in 100. We are justified in rejecting the hypothesis that there is no difference between the groups.

TABLE 2

FAVORABLE AND UNFAVORABLE MENTIONS BY USERS AND NON-USERS OF INSTANT COFFEE
(REVISION OF HAIRE'S TABLE 1)

The woman who buys Nescafé is seen as:	*By Women Who Had Instant Coffee in the House (N = 32)*			*By Women Who Did Not Have Instant Coffee in the House (N = 18)*		
	Number	*Per Cent of Subjects*	*Per Cent of Mentions*	*Number*	*Per Cent of Subjects*	*Per Cent of Mentions*
Economical	22	69	(39)	5	28	(15)
Good housewife, plans well, cares about family, plans balanced meals	18	56	(33)	2	11	(6)
Total Favorable Mentions	40	125	(72)	7	39	(21)
Not economical	0	0	(0)	2	11	(6)
Lazy	6	19	(10.5)	7	39	(21)
Poor housewife, does not plan well, does not care about family	10	31	(17.5)	17	95	(52)
Total Unfavorable Mentions	16	50	(28)	26	145	(79)
Total Mentions	56		(100%)	33		(100%)

Because Haire's objective was to reveal the presence of underlying attitudes toward instant coffee, Table 2's division of responses into favorable and unfavorable categories seems better suited for analysis. Note particularly the bracketed percentages listed under "Per Cent of Mentions." Rather than base respondent mentions upon the N of respondents, the new base is taken as the total mentions made by both users and non-users. In this manner it is possible to compare relationships between the groups. As Table 2 indicates, they are not as dramatic as those in the original presentation.

The average number of mentions by non-users (33 by 18) and by users of instant coffee (56 by 32) is roughly equal at 1.8 per respondent. In order to assess that statistic, however, we need to know the ranges as well. That is, did a small but vocal minority contribute the bulk of the mentions to the detriment of a larger but mute group? Such a statistic would provide clues for an understanding of the attitude classes in Table 1's stub.

Bias By Baking Powder?

Refer back to the grocery list used. Note that Rumford's Baking Powder immediately precedes the coffee item. Is it possible that baking powder, as a symbol of the good cook, could have subtly biased the responses? It may have exerted a strong "contrast effect" so that instant coffee assumed a shameful role in obvious contradiction to the stereotype of what constitutes a good cook.

An analysis of respondents' verbatim reports published by Haire augments that suspicion. That is, in the ambiguous evaluative situation, respondents seem to have searched the list for items to serve as guides. For example, one respondent resolved her ambiguity in this fashion:

". . . she must like to cook and bake as she included baking powder. . . ."

Another rationalized baking powder in terms of other items:

". . . canned peaches, instant coffee . . . all the items with the exception of Rumford's are easily prepared items."

It seems safe to conclude that baking powder did prejudice the case against instant coffee, but it is impossible to know to what extent.

Westfall, Boyd and Campbell Study

This second study of attitudes toward instant coffee was conducted seven years later by the above Northwestern University researchers. After first qualifying respondents as to recency of coffee-making, they asked:

". . . the last time you made coffee—did you use percolator, drip method, vacuum method, instant coffee, or some other method?"

Next they asked respondents to match 19 personality dimensions (the best cook you know, the worst cook you know, etc.) to each of the coffee preparation methods. This technique, the authors suggested, would reveal underlying attitudes toward instant coffee.

Generally, their findings agree in direction but not degree with Haire's. For example, "dislikes to cook," "thrifty," "cares for family" and "good planner" proved to be in good agreement. But Westfall's "poor cook" dimension had smaller

differences between users and non-users than did Haire's. Surprisingly, "laziness" proved to be similar for users and non-users of instant coffee in Westfall's findings.

This last finding, contradicting as it does expectations that users and non-users should react differently to "lazy" attitudes, suggests a common weakness of forced choice research techniques. There's simply no way of knowing to what degree forcing elicits rationalizations rather than true attitudes. And unless attitude intensity is measured, superficial responses must be accepted as equally valid as expressions of rigid attitudes. Consequently, the "laziness" dimension may be as much a rationalization as "dislike for flavor."

Force of Habit

Despite projective researchers' apparent ability to reveal underlying attitudes that are ostensibly more sophisticated than those elicited by direct questioning, there is no basis upon which to evaluate the worth of those "revealed" attitudes. In several reported studies, different researchers probing for underlying attitudes about identical products have unearthed disparate findings.

Therefore the suspicion emerges that these so-called underlying attitudes may be rationalizations also—but at a slightly more sophisticated level; that is, they may be less ready rationalizations about cultural stereotypes that do not necessarily impinge on decisions to buy or not to buy a given product. What's more, these underlying attitudes may be rationalizations for habitual or routinized behavior.

Pursuing this notion of habit, we know that psychology has well documented the fact that resistance to change usually increases with age. Therefore, in order for advertising researchers to assess properly opportunities for new products, or to understand their share-of-the-market for established products, they should consider seriously, much more seriously, the workings of neophobia (fear of the new) as a sales deterrent. This assumes sharper meaning if we remember that one objective of advertising established impulse items is to reinforce conditioned brand responses. Therefore, research for any brand in a given product line needs to measure as accurately as possible the drag of previous conditioning and how that drag is distributed among competitors market by market.

This is to suggest, then, that Haire and Westfall failed to tap anything of significance so far as "true attitudes" toward instant coffee are concerned. Both studies made it easy for respondents to report superficial attitudes that may actually represent the weight of habit, not subconscious attitudes as implied. It is hoped that this suggestion casts the concept of depth research into a new focus.

REFERENCES

HAIRE, MASON. Projective Techniques in Marketing Research. *Journal of Marketing*, Vol. 14, No. 5, April 1950, pp. 649-656.

SELLTIZ, CLAIRE, MARIE JOHODA, MORTON DEUTSCH AND STUART W. COOK. *Research Methods in Social Relations*. New York: Henry Holt & Company, 1959.

WESTFALL, RALPH L., HARPER W. BOYD AND DONALD T. CAMPBELL. The Use of Structured Techniques in Motivation Research. *Journal of Marketing*, Vol. 22, No. 2, October 1957, pp. 134-139.

"Sir—In your issue of December 31 you quote Mr. B. S. Morris as saying that many people are disturbed because about half the children in the country are below average in reading ability. This is only one of many similarly disturbing facts. About half the church steeples in the country are below average height, about half our coal scuttles below average capacity, and about half our babies below average weight. The only remedy would seem to be to repeal the law of averages. Yours faithfully, etc."

—LETTER TO THE EDITOR OF THE *Times*, LONDON, JANUARY 1954.

Letters

Comment on "How Incentives and Subsamples Affect the Precision of Mail Surveys"

JACK B. HASKINS
The Curtis Publishing Company

THIS JOURNAL "is intended for practitioners and users of advertising research" and from this it may be inferred that the papers should therefore be practical and usable. Also, "reports of findings are favored over theoretical discussion." While this should not be narrowly interpreted, theoretical discussion should be translated into terms that have some relevance to the work of advertising researchers.

The subject article (published in this *Journal*, Vol. 1, No. 1, September 1960) can be criticized on both those grounds as well as on a more serious one, the validity of the findings.

1. Under the conditions of the two surveys reported, it is impossible to make any valid comparison of relative costs or relative effectiveness of the techniques involved.
2. Some rather specialized usages of such terms as "precision," "information" and "effective response rate" are put forth which, if not irrelevant to the topic at hand, are ineffective as communications to the reader of the article.
3. In general, the article is not clear on the answer to the question, "How do incentives and subsamples affect the precision of mail surveys?"

Each of these points is discussed in more detail below.

Validity

One stated objective is "to compare the two procedures in terms of the amount of information obtained per dollar spent to collect it." In order to make such a comparison, it would be necessary to hold all other factors constant except the financial incentive, under the present design; this would not, of course, be necessary in the case of some other design, such as a factorial experiment, where several factors could be tested at once. In this study, there are several other things which varied and which confound the results in terms of the stated objective.

a. The studies were conducted in different years. The composition of the population had changed during the four years between studies, along with possible other broad social changes. These differences could conceivably produce differing response rates to mail questionnaires, independently of the incentives used.
b. The questionnaires were different, although judged by the investigator to be "comparable in length and complexity. . . . Some topics were added in the 1960 study and it was decided to split the sample and ask certain questions of half the sample only." Differences in questions asked could certainly influence the response rate.
c. An advance postal card was sent to the 1960 sample. Methodological studies by The Curtis Publishing Company demonstrate that an advance postal card can increase the number of returns (as well as the quality of returns in terms of predictive value).

This, then, is a comparison of results from two studies which differ not only in respect to the financial incentive, but from at least three other factors which confound the results so that no definite comparative statement can be validly made about the effects of the incentive, nor about the costs attributable to it. This, incidentally, is the Achilles heel of another popular statistical maneuver, the multiple regression analysis of historical data, in which no amount of statistical manipulation can overcome the confounding of variables in the original data.

Semantics

Several terms are used which appear to be either inappropriate or poorly defined.

"Effective response rate" is defined as "the percentage who replied to the first mailing, two times the percentage who responded to the second mailing, and four times the percentage who responded to the third mailing." Actually, this is the "predicted response rate if all non-returnees had received additional mailings." (Incidentally, the article shows 1956 as being somewhat better than 1960 on this statistic, whereas 1960 was actually somewhat better than 1956 when the arithmetic is corrected.)

"Precision" is used frequently but not defined. According to one source (Palmer O. Johnson, *Statistical Methods in Research,* New York: Prentice-Hall, Inc., 1949), precision is the amount of information which a particular experiment affords. "Efficient" is also used but not defined; Johnson defines an efficient estimate as one whose sampling distribution tends to the normal law with the least possible standard error as the number of observations is increased. Just how relevant this statistical kind of "precision" and "efficiency" is to the topic at hand is not stated. In any event, they are somewhat removed from the working language and requirements of advertising researchers; the same can be said of the term "amount of information" which is defined as the reciprocal of the variance of an estimate. This is not to say that these terms are irrelevant, but that their relevance is not made clear.

Relative Cost

The statement is made that "At first sight the collection costs involved in the 1960 study would seem to be much greater than in the 1956 study." But the 1956 study produced 5,288 responses at a cost of $11,980 or $2.27 per returned questionnaire; the 1960 study produced 6,347 responses at a cost of $12,471 or $1.96 per returned questionnaire. To produce a given number of responses, the 1960 study is obviously cheaper.

Conclusion

Apart from these deficiencies, one other scientific principle should be considered and that is the law of parsimony, which is to say that one should choose the simplest method possible to describe a given phenomenon. In an applied field such as this, and in this *Journal,* parsimony is even more important than in the academic field.

Rejoinder

Lester R. Frankel
Audits and Surveys Company

The points raised in the comment on my article can be grouped into five general classes. These are:

1. The article is too theoretical.
2. The article is neither practical nor usable.
3. The development of the thesis is invalid.
4. The article is deficient on technical grounds such as, for example, incorrect usage of statistical terms and an error in arithmetic.
5. Miscellaneous points, such as (a) the article attempts to use statistical manipulations similar to those used in multiple regression analysis of historical data to overcome the confounding of variables and (b) the selection of the article for publication by the editors violates "the law *(sic)* of parsimony."

I shall discuss each of the first four points and ignore the fifth.

Point 1

The critic is explicitly concerned about the article because of its use of theory. In addition, much of his criticism appears to be dominated by a reluctance to accept theory as a basis for practical guidance. Without the acceptance of theory an experiment could only be conducted for its own sake. The best that could be achieved with this philosophy in a mail questionnaire experiment to test alternative techniques would be to use trial and error tests and come up with a decision as to what *would have been* the best technique. Without a theory the utilization of even these results at a different time or with a different questionnaire would be impossible because the conditions would not be identical.

Point 2

It is claimed that the article is neither practical nor usable. Theory is always abstract; but the test of good theory is its eventual application. By this

standard the article certainly passes the test. From our theorizing we were able to get more information (as defined by sampling statisticians) for our money. Surely this is practical.

Point 3

It is claimed that the method used to develop the thesis of the article is invalid. The specific criticism against the approach is that the 1956 and the 1960 studies cannot be compared because of (1) difference in time, (2) different questionnaires and (3) the use of preliminary post cards in the 1960 but not in the 1956 study. Naturally, the post card variable was not present in 1956. This, together with the 25-cent incentive in the first mailing, was introduced *by design* into the 1960 study. The fact was clearly stated in the article; first in the paragraph of the text of the article which described the incentive program, and second in Table 2 when the cost of the post card mailing was included with the incentive costs. It appears that the original statements in the article were deliberately misconstrued to form a basis of an attack. To becloud the issue still more the critic has on his own initiative introduced the word "financial" prior to incentive to give the impression that post cards were not part of the incentive program.

It is reasonable to assume that the two studies were comparable in spite of the differences in time and in the questionnaires. In both studies three questionnaire mailings were made; a five-dollar cash incentive was used in the third mailing and almost identical over-all sample representations were obtained. The effect of the added incentives in 1960 was to accelerate the responses among those persons who would eventually respond.

Point 4

It is claimed that the article has many technical shortcomings such as the incorrect usage of certain statistical terms and an error in arithmetic. As a field develops a technical vocabulary evolves. Its terms become symbols for ideas, concepts and operations. When a new field develops from an old one, it is natural that many of the earlier terms should be carried over. So it is with sample survey methods which have developed from statistical theory. Many terms retain an identical meaning; others are used in closely similar senses. The exact meanings of its terms can, of course, easily be found in books on sampling, as for example: Hansen, Hurwitz and Madow, *Sample Survey Methods and Theory* and Deming, *Some Theory of Sampling*. In writing an article on some aspect of survey sampling the author would naturally use the terminology of survey sampling. If certain of the terms are in general usage by potential readers of the article these terms are not defined. If other terms are used to a lesser extent, or if certain new terms are introduced, definitions are given.

In this area my use of the terms "precision" and "efficient" are questioned. It is claimed that neither term was defined (actually the word "efficiency" was defined) and that the relevance of these terms is not established. The critic then attempts to define these terms. However, instead of obtaining these definitions from a standard text on sampling he refers to a book on statistical methods. It happens that the word "precision" has almost the identical meaning in sampling as it does in the definition which is cited. If the critic concedes that a mail questionnaire survey is, in effect, an experiment, then the use of the term "precision" under his definition is relevant. The term "precision" in books on sampling may have reference to any sampling process where probability methods of selection are used.

The danger of going outside of a particular field of application for definitions is illustrated when texts on sampling are bypassed by the critic in seeking the meaning of the word "efficient." He misses the point when he cites a definition of an "efficient *estimate*." This definition, due to R. A. Fisher and widely accepted by statisticians, refers to a criterion for selecting a formula to predict population parameters from corresponding sample statistics. In this sense of the word, the "efficient" formula is the one which predicts the parameter with least variance for a fixed sample size. For example, in most cases the arithmetic mean is to be preferred over the median as an estimate of the mean of the population. No wonder Mr. Haskins has difficulty in seeing the relevancy of this concept to the operation of mail surveys.

All of the computations have been rechecked and no arithmetic errors were found.

* * *

I have no fear that advertising researchers will encounter conceptual or technical difficulties in reading the article. Most advertising researchers are versed in the scientific method; they have a healthy respect for theory; they are familiar with current sampling terminology; they are able to reason correctly; and finally, they are willing to learn from the experience of others.

FEDERAL STATISTICS IN ADVERTISING

INGRID C. KILDEGAARD
ARF Research Statistician

THE FEDERAL CENSUS OF 1870 was the first to use a tallying machine—data from the previous eight censuses were tabulated by hand. Twenty years later electrical tabulation machines were introduced, but they were a far cry from 1960's data processing equipment. Today, with the help of high speed computers and printers, the Bureau of the Census is producing the results of the decennial census more rapidly than ever before. This added speed will enhance the value of census data to those who want to know more about their markets.

Small Area Statistics

Especially valuable will be the tabulations of data from small geographic areas. In previous years, most of such information was out of date by the time it was offered. Available while still current, these small area statistics will be useful to describe the people that compose the local market.

Different neighborhoods within a city obviously have different market potentials. The advertiser who wants to reach a specific segment of the market will find it profitable to compare living conditions, economic status or other characteristics of these areas. He could even combine them to correspond with his company's own sales districts. A map showing actual sales by district, and those population or housing characteristics relevant to sales, would be a valuable device for evaluating sales performance, setting sales quotas or planning product distribution.

More Efficient Samples

With the increased use of consumer or household surveys, it is important to the advertiser that his samples be planned efficiently and economically. Here the survey designer will find small area statistics a valuable resource. Suppose he wants to measure the penetration of an advertising campaign in a test city. He selects a sample, first using the census data to stratify blocks according to population characteristics. He numbers the blocks on a city map and selects a prescribed group from each stratum. He then designates the households (or perhaps a subsample of the households) in the sample blocks for the interview. With the materials available from the Census Bureau, designing an efficient sample is relatively simple and inexpensive.

Here are some of the areas for which these materials will be furnished.

Census Tracts

For statistical purposes, large cities have been arbitrarily divided into small geographic areas. At present there are 23,000 such census tracts, nearly twice as many as in 1950. The 175 places "tracted," each with one or more cities, include most U. S. cities over 50,000 in population. The average tract has a population of about 4,000. Local committees set boundaries that make tracts homogeneous in population and housing characteristics. Boundaries are permanent: even if a tract is subdivided the resulting smaller tracts are still contained in the original. Thus changes over time can be analyzed.

Tract reports scheduled for May through September 1961 will include population by age, education, color, sex, marital status and relationship, occupation, family income, means of transportation to work, average rent, property value, number of rooms, persons per room, condition of plumbing, etc. Photostats of advance data are now available at $3.50 per 100 tracts.

City Blocks

These statistics will be published in a series of about 500 reports from March through August 1961. Population and housing data, reported in some of the same ways as census tracts, will be shown by city blocks for each city where the last official census count was 50,000 or more, and for about 250 places included by special request. Maps will be provided on which the blocks can be located.

Enumeration Districts

"ED's" are the small areas determined by the Bureau especially for the decennial censuses. In 1960 there were 275,000. Their populations range from a few people to two or three thousand, and average about 700. Data are not published for individual enumeration districts, but anyone may obtain at nominal cost photocopies of final population and housing counts along with maps showing district boundaries. Some time in 1961, summary cards or tapes of the ED data will be prepared and offered at cost.

Further information and cost estimates are available from the Director, Bureau of the Census, Department of Commerce, Washington 25, D. C.

RESEARCH IN REVIEW

GWYN COLLINS, THORNTON C. LOCKWOOD and GEORGE M. SHIREY
ARF Research Associates

Sampling Made Palatable

W. Edwards Deming. *Sample Design in Business Research.* New York: John Wiley & Sons, Inc., 1960, $12.00.

This book is of major importance to advertising researchers. Although Professor Deming maintains in his preface that it is a textbook written from the viewpoint of a theoretical statistician it is a more lively, more interesting and more individual book than his description suggests.

Divided into three unequal parts, it is at once a manifesto, a handbook and a text. A memorable introduction deals with the statistician's function as Professor Deming sees it. He describes the responsibilities and duties of the statistician to his client or employer in a chapter which is a statistician's bill of rights. Probably the consulting statistician will find it easier to operate within Professor Deming's rules of operation than will the salaried statistician. This chapter can be read with profit by all who have and all who give statistical help.

The remainder of the first part of the book is devoted to an eminently sensible discussion of the materials with which sampling statisticians must work. This constitutes a reversal of the method of many texts on applied statistics which first develop the theory and introduce the reader to the facts of life later. Professor Deming first lays a foundation of common sense on which to build later and more abstract ideas.

The second part of the book introduces replicated sampling. Basically the method consists of collecting data from several subsamples and treating the subsamples separately instead of working with one large sample. The procedure has become increasingly popular, and the author mentions that he has used it exclusively since 1948. This book offers the most detailed description of the method and its advantages currently available. Even those who are easily discouraged by simple statistical theory will find Professor Deming's discussion of this method easily understood and the method attractive. The procedure is illustrated in great detail with some well chosen examples.

The third part of the book is theoretical and explains many of the ideas which have been introduced in earlier sections. Sampling is an applied science and as such it must draw on many results and theories derived from probability and statistical inference. In the brief space available to him, Professor Deming has therefore had to take a good deal for granted. Moreover, he has not aimed at covering every point of the relevant statistical theory, but has concentrated his energy on a few important topics. After developing the idea of expected value and some theory of variances he has a note on some distributions of rare events and a method for transforming them into a more familiar form. He has also a brief and lucid account of the theory underlying the best selection of strata for stratified sampling.

This partial listing of the contents of this book may be misleading. It is not a conventional sampling cookbook. Professor Deming is one of the most widely known consulting sampling statisticians, and his experience in this field has been lengthy and diverse. In this book he shares much of it with the reader. It is in some ways a professional autobiography and includes much illustrative material of sampling practices, both good and bad. The author has made this a better book by exposing something of himself in this unlikely vehicle, a statistical text.

It should be said too that his account of sampling practices is written from a definite point of view: that of replicated sampling. There are other sampling methods but for them the reader will have to turn to such standard texts as *Sample Survey Methods and Theory* by Hansen, Hurwitz and Madow.

Something should be said about the mathematics the reader needs in order to understand this book. The first part of the book is completely unmathematical. The second part can be read with only high school mathematics. The third part makes no great extra demands, but will be read more easily by those with rather more algebra and some calculus.

About half a dozen well known texts currently delineate the field of sampling statistics. This is a significant addition to them. It should be on and off every researcher's book shelf.

—GWYN COLLINS

Persuasion on the Procrustean Bed

Herbert I. Abelson. *Persuasion: How Opinions and Attitudes Are Changed.* New York: Springer Publishing Co., Inc., 1959, $3.75.

Procrustes, you may recall, was the legendary robber who stretched or trimmed his victims to fit upon an iron bed. The critic who decries the disorganization of research findings is usually the first to accuse anyone who tries to organize them of using Procrustean techniques. In the face of such inevitable criticism the author is to be congratulated on this attempt to provide a framework for exploring the psychological bases of persuasion.

Dr. Abelson has drawn extensively upon the literature of the social sciences to reach 24 conclusions grouped in five chapters: how to present the issues, the influence of groups, the persistence of opinion change, the audience as individuals and the persuader. His purpose is to weigh the factors responsible for the success or failure of attempts to persuade people.

The advertising researcher is likely to be intrigued at the prospect of such an undertaking but disappointed before completing the first chapter. Conclusion number two states that "When opposite views are presented one after another, the one presented last will probably be more effective." It is supported by results of four experiments. *The Order of Presentation in Persuasion,* edited by Carl I. Hovland, and one of the 139 entries in Dr. Abelson's bibliography, states this principle somewhat differently: "When two sides of an issue are presented successively by different communicators, the side presented first does not necessarily have the advantage." Dr. Hovland's conclusion is followed by results of eight experiments in which the relative effectiveness of first and last communications varied widely depending on the experimental conditions. None of these eight studies is cited by Dr. Abelson.

The author does recognize the danger of interpreting such research data too broadly, warning that slightly different experimental conditions might produce different results. He properly notes that his material "should be regarded by the reader as something to compare with his own experience, rather than as a substitute for experience." The reader will be struck nevertheless by the use of fairly pat answers when highly qualified interpretations seem more in order. The author only partially succeeds in reducing this impression in the discussion section which follows each conclusion and its research support.

The detailed descriptions of 62 specimen studies are of considerable value, but there is one important omission. The people studied are identified as students in 25 studies, as soldiers in six, but in 20 cases are not described at all.

Although the preface holds out some hope that opinion and attitude change data relevant to the advertiser will be included, only four of the studies can be considered to concern product-oriented attitude change. Foreign policy, domestic government and non-government issues, and various issues concerning attitudes of small groups dominate the topics cited in support of attitude change.

We may ask, at this stage of development in the science of persuasive arts, whether summaries should be attempted at all. The answer must be an emphatic "yes" if only because they, like this volume, present such a wealth of detailed and thought-provoking source material. Such organization of sources may be the most persuasive argument of all for more systematic research.

—THORNTON C. LOCKWOOD

Motivation Research: Lest We Forget

Lawrence C. Lockley. *Use of Motivation Research in Marketing.* New York: National Industrial Conference Board Inc., 1960, $1.50 to NICB associates, $7.50 to non-associates.

At the outset the author points out that enough time has now elapsed to permit an appraisal of the techniques used by the motivation researcher. Specifically, his purpose in preparing this report was to describe these techniques, appraise their value, and indicate where and how they can be used to greatest advantage.

More than a third of the monograph is devoted to descriptions of such techniques as the Thematic Apperception Test (TAT), paired pictures, sentence completion, word association and depth interviews. Each is illustrated by a practical application, although none is very detailed.

Following this review is a section on suitable applications of motivation research. Dr. Lockley

points out that motivation research is ideal for doing the preliminary work necessary to design a large-scale study using more conventional survey methods. However, when the questions asked do not induce rationalizations, or when there are apparently no important associations in the preconscious or unconscious mind with the product or service being studied, he concludes that psychological techniques serve no purpose.

In his conclusion, Dr. Lockley notes that motivation research itself is new only in its application to the study of why people react as they do to various marketing stimuli, and its success in this area has not yet been established. Important, and certainly stressed by Dr. Lockley, is the fact that very special skills are required to design and use psychological techniques in marketing and to interpret the data which result.

The author suggests that the most appropriate uses of psychological motivation research are in developing ideas, in searching consumers' minds for motives that can be proved or disproved by further research, and perhaps in planning and pretesting advertising copy.

Dr. Lockley's paper is recommended for anyone who has not had the opportunity to read any of the published material on motivation research and who would like a basic outline of this approach. But it adds little that is new to the published literature of motivation research.

—GEORGE M. SHIREY

Attitude Scaling at Loose Ends

F. P. Kilpatrick and Hadley Cantril. Self-Anchoring Scaling, A Measure of Individuals' Unique Reality Worlds. *Journal of Individual Psychology*, Vol. 16, No. 2, November 1960.

"A self-anchoring scale is simply one in which each respondent is asked to describe, in terms of *his own perceptions, goals and values*, the top and bottom, or anchoring points, of the dimension on which scale measurement is desired, and then to employ this *self-defined* continuum as a measuring device."

The authors go on to maintain that it makes sense to compare respondents on the basis of goal-related variables with psychological meaning (such as striving, where one is now and will be a few years hence), but only if the individual's "unique reality world" is taken into consideration.

The method consists of such things as open-ended probing, content analysis and non-verbal scaling. First, the respondent is asked to consider a dimension such as "way of life," and then describe the best possible example of such a dimension. Probing is continued until the respondent can contribute nothing further. A description of the worst possible example of the dimension is similarly obtained. Second, the respondent is handed a pictorial ten-point ladder scale and told that the best and worst examples he has just described should be regarded as end points on the scale. Usually the respondent is asked, "Where on this ladder would you say you are now?" Past evaluations and future expectations can be obtained using the same scale. A third step may involve probing for reasons for different scale positions.

Can the self-anchoring of one respondent be compared with the self-anchoring of another? The authors say yes: "The scale level selected by one person or group (average of selections) can be specifically and meaningfully said to be higher, lower or equal to the scale level of some other individual or group, because the frames of reference of the replies are in fact similar psychologically." The article continues with comparisons of average rating scores on a personal "way of life" scale for each of five groups of adult Americans; on a national "aspirations and fears" scale for each of seven groups of foreign parliamentarians; on personal and national "aspirations and fears" scales for Indian and American students; and on personal and national "aspirations and fears" scales for South African primitives.

In all of the foregoing the authors seem to have missed an important point. A scale is simply a device for generating numbers, usually one for each respondent. The problem is how to use these numbers. Unfortunately, the authors apply this scaling device to "way of life" and "aspirations and fears" dimensions which appear to have no behavioral determinants in the real world. The authors' maintenance of "psychological similarity" to justify averaging the scale levels of respondents highlights an undesirable influence in psychology: too much emphasis on what the psychologist thinks he is doing and too little on what the respondent is doing. Advertising researchers who use or interpret this device will be well advised to seek comparisons of the resulting scale numbers with actual behavior.

Many readers would have been interested in an analysis of the dispersion of respondents' self-ratings on the self-anchoring scale compared with

those obtained on a fixed scale such as the semantic differential. However, it would still remain to be proved that any differences obtained reflected differences in what was measured rather than simply the effects of different techniques.

—THORNTON C. LOCKWOOD

How Not to Lie with Statistics

Joseph Alsop. The Wayward Press: Dissection of a Poll. *The New Yorker,* September 24, 1960, pp. 170-184.

Bernard D. Nossiter. You've Never Been Quizzed by a Pollster? *The Washington Post,* October 16, 1960, p. E3.

Advertising and market research usually precedes some executive action. To verify the research is to check the stable doors after the horse has bolted. Often it throws good money after bad to make such checks and they are rarely called for. The Presidential political polls are different. They provide, once every four years, a great public examination of the state of the survey art.

Since 1948 the pollsters have been beset by unbelievers, all ready to cry foul. Mostly they have been the enthusiastic supporters of some candidate who believes himself hurt by the results. Some have warned that published polls influence the vote. A minority of the critics consists of those who would, but cannot, see how the final poll figures are derived. For this minority, these excellent articles cast light into the dark places of polling methods.

Mr. Alsop's article may have been inspired by the surprising Gallup Poll of August 17 which found 50 per cent of the electorate favoring Mr. Nixon, 44 per cent favoring Mr. Kennedy, and only six per cent undecided. Everyone was disgruntled; Mr. Kennedy because he started too low, Mr. Nixon because he started too high, and the rest of us because the six per cent figure seemed incredibly small. In tracing these figures to their source Mr. Alsop brings the spotlight to focus, in turn, on Dr. Gallup's sampling methods, his weighting and adjusting processes, his record and his influence.

Mr. Nossiter's article makes splendid collateral reading. His is a detailed and illustrated account of how Dr. Gallup "deep fries, boils, roasts and sautés a mere 1,600 interviews." Interestingly, Mr. Nossiter concentrates on the poll of September 25 at which time, he says, Dr. Gallup's method had been amended, apparently at Mr. Alsop's suggestion.

Up to now the publicity given to polling methods can best be described as creeping disclosure, or perhaps progressive revelation. Certainly, it has never been easy for just any member of the public to find out where the published figures came from. This must be far from Dr. Gallup's wish. He says that he has never refused to answer any question about his methods and that he has supplied editors who subscribe to his service full technical information. If this is so, then we must assume editors, naturally reluctant to serve up this less sensational aspect of polling, have not experienced much call for it. But dormant suspicions have now fed on so many polls, estimates and guesses that the time has come for explanations. Here, in the articles under review, are probably two of the best.

Alsop and Nossiter give a good explanation of one of Dr. Gallup's most misunderstood adjustments. Readers of this *Journal* will sympathize with his problem in having to predict not just the state of opinion generally but that of those people who are going to vote. Moreover, he must rely on good fortune to have the vote of the electoral college fall in the same direction or he runs the risk of more public outcry. So Dr. Gallup throws out of his sample respondents he judges will not vote. These include aliens, people who say they do not intend to register, and those who say they don't intend to vote. These are more often Democrats than Republicans. It may not be the best way of sorting out the non-voters, and it certainly is not the only way, but it is defensible and Dr. Gallup does defend it.

Both articles explain, too, the small numbers of undecided voters the poll finds. Dr. Gallup asks those respondents who describe themselves as uncertain which way they lean and in this way disposes of all but about six per cent who refuse even to lean.

These two points are but examples of those touched upon in these articles. Not only do they answer those questions we were all asking recently, but they examine, too, why these questions were asked and why answers were so long delayed. While Mr. Alsop complains of Dr. Gallup's habit of "mystification," Mr. Nossiter reports that Dr. Gallup complains of editors. Dr. Gallup maintains that he tries to tell all in six languages, in the technical journals and in popular descriptions. In addition, he says he thinks "basic" changes in technique should be reported.

Messrs. Alsop and Nossiter agree that the pollsters are honest men, an opinion we all can hardly fail to share. But it is evident that in these hard times one must not merely be honest but must readily appear to be so. In this latter effort Dr.

Gallup is apparently impeded by some of his subscribing editors. No blame can be placed on an illiterate or innumerate public. Mr. Alsop has shown how to make technical information more than just palatable in a popular magazine, and Mr. Nossiter has done so in a major newspaper. Let us hope that in 1964 all pollsters' editors will profit from their example. —GWYN COLLINS

Two for Research: Schramm & Alderson

C. H. Sandage and Vernon Fryburger (Editors). *The Role of Advertising: A Book of Readings.* Homewood, Ill.: Richard D. Irwin, Inc., 1960, $8.00.

Always valuable to the student is convenient access to a wide variety of papers relating to his subject. Collections of such material may be even more valuable to the "student" who has completed his formal education.

With this in mind the editors have compiled a book of 49 papers by 42 individuals about five roles of advertising: in the economy, as communication, in society, in business and in social responsibility. Their purpose was to foster the development of a philosophy of advertising by bringing together provocative statements of those who have contributed to its rational explanation.

Although Messrs. Sandage and Fryburger would surely agree on the importance of research in understanding advertising, the book does not dwell long on this facet. Research is represented by only two papers. Fortunately, they are both excellent: "The Effects of Mass Communications" by Wilbur Schramm and "The Productivity of Advertising Dollars" by Wroe Alderson.

Dr. Schramm reviews mass communications research since its birth some 25 years ago. Lazarsfeld's early voting studies, his studies of personal influence, later investigations carried out in psychological warfare, and opinion research in the armed forces and in advertising are given as evidence of what mass communication can and cannot do. We soon realize that it is severely limited in its capacity to change opinions, values or behavior. Research thus contradicts the advocates, well-represented in this book, of the position that advertising has become a major wielder of social control in America and that many are using this power irresponsibly.

In contrast to Dr. Schramm's paper, Wroe Alderson presents theoretical approaches to two basic questions: How much should a firm spend for advertising? Given an advertising budget of a stated size, how should the money be spent?

Based on the principle that advertising is subject to the law of diminishing returns, Mr. Alderson shows how to develop an advertising productivity curve which can be used to determine the optimum size of a budget. Although the data often may be faulty and many problems exist in implementing proposed models, Mr. Alderson points out that these can be used to form a basis for relative judgment and a move in the right direction.

The Role of Advertising contains many stimulating papers. A few are classics; many are familiar to the practitioner. Together they provide a useful supplement to standard texts.

—GEORGE M. SHIREY

If You Can't Audit, Watch the Purchase

Earl E. Houseman and Benjamin Lipstein. Observation and Audit Techniques for Measuring Retail Sales. *Agricultural Economics*, Vol. 12, No. 3, July 1960, pp. 61-72.

Retail audits have long been a standard procedure for evaluating new consumer products and checking the progress of older ones. Although alternative and less expensive techniques have recently been much discussed, it is a safe bet that store audits will be with us a long time yet.

The advantages of the method are obvious. A store inventory check, a record of dealers' purchases, another inventory, a little arithmetic and we have sales. Each week, or month, or two months, regularly as clockwork the figures roll in. Unfortunately, not all the figures roll in; for some of the biggest stores refuse to cooperate though researchers may plead and cajole.

From the ashes of these endeavors has arisen a new technique for getting the missing figures: that of actually observing consumer purchases within the store. In this paper Messrs. Houseman and Lipstein explore the best way of doing it. They describe the measurement by the two techniques, observation and audit, of the sales of apples and pears in the produce department of 60 stores in Philadelphia. The two methods are compared for the precision with which they estimated the week-to-week changes in sales.

The authors point out that the two methods are not completely comparable in this way. Each provides information which the other does not. For

example, the audit procedure gives inventory levels while the observation technique provides details of the number and nature of consumer purchases.

When retail sales are measured in a city or some larger area the audit method usually uses only a sample of the stores in the area. Estimates of area sales will, of course, contain a sampling error because the sample stores do not always exactly reflect the behavior of the other stores not in the sample. The observation method, too, is subject to this same sampling error and more, since usually only a sample of each store's business hours are observed. Generally, then, we expect the observation technique to be less precise than the audit technique. But the cost of the two methods is not the same and the question arises, not which method is most precise, but which offers most precision for some fixed amount of money.

The authors' conclusions are quickly stated: the observation method is better for fast-selling products in large stores, and the audit method is better for slow-moving products or small stores. But there is more in this paper than just that. The methods considered for obtaining the best observation estimate—varying the length of the observation period; stratifying the hours available for observation by degree of consumer activity; and estimating product sales as a ratio of purchases of the product to purchases of all products—all these will help practitioners. The description of the experiment is clear and will be enjoyed by researchers of statistical bent.

The significance of the paper is perhaps greater now than when this study was started in 1955. More and more stores refuse to cooperate in store audits and, unfortunately, these are the biggest and most important for sales estimates. Such stores not only have overwhelming sales volume but their sales pattern is often different, with private label brands playing a big part.

Here Mr. Houseman and Mr. Lipstein have studied the perfect supplement to the audit. It flourishes in big stores where the audit has foundered. It thrives in busy periods which are anathema to audits. It adds a valuable dimension of information on the size of consumer purchases.

This study brings into focus many of the promises and problems of the observation technique. The reader will also be grateful for the thoughtful notes added to the paper by Lester Frankel and Solomon Dutka. They provide an excellent summary of the special utility of several methods of measuring retail sales.

—GWYN COLLINS

The Short Happy Lives of TV Commercials

Schwerin Research Corporation. When Should the Effects of Television Advertising Be Measured? Part I: Recall. *Technical and Analytical Review*, No. 4, Spring 1960, pp. 1-16.

The answer to the question posed by this pamphlet seems to be "periodically."

Written at a non-technical level, it first examines selected learning theory experiments by Luh and Kreuger, who found learning and retention to vary depending on the task to be learned, measurements made and time intervals used. Schwerin warns that the motivation variable—positive in Luh's and Kreuger's test environment, but of unknown influence in the advertising environment—will also serve to restrict the generality of psychological learning theory data to advertising.

The theme of the pamphlet may be summarized as follows: what is the effect of television commercial repetition upon learning? The criterion for learning is recall of a commercial at different points in time.

Three measures of recall are used: total responses (brand name and 20 pre-coded ideas) per 100 persons, volunteered recall of slogan, and share of audience recalling brand name and at least one sales point. This interesting method highlights the largely subjective nature of recall measures in general.

Schwerin maintains that changes in the level of recall after two or four weeks suggest positive or negative motivation leading to the learning or forgetting of a commercial. The reader will accept this conclusion more readily if he believes that the initial recall measure was free of any group-inspired motivation associated with the theater technique, and that one sample of respondents can be compared legitimately with another.

The pamphlet does not analyze results statistically to determine whether some levels of obtained recall differed significantly from others. This may reflect the authors' desire to simplify results or to recognize that the samples are not comparable. The latter condition may invalidate the comparisons made in this study, but those individuals responsible for gauging the memorability of TV commercials should value such comparisons. They rarely appear in the published literature.

Readers who question whether commercial recall is an appropriate measure of television's effect will be interested in another Schwerin pamphlet to be published soon. It promises to report television's effect on changes in attitude and behavior.

—THORNTON C. LOCKWOOD

The Anatomy of New Product Success

Louis A. Fourt and Joseph W. Woodlock. Early Prediction of Market Success for New Grocery Products. *Journal of Marketing*, Vol. 25, No. 2, April 1960, pp. 31-38.

In this interesting paper Messrs. Fourt and Woodlock describe their method of predicting the market success of new consumer products. Although the approach they describe can be applied equally well to products new or old, they point out in their first paragraph that four out of five new products launched between 1945 and 1955 are reported to have failed. This emphasis on new products is further justified by another opinion, quoted by the authors, that it usually takes 19 to 24 months to know whether a new product is successful or not.

Now the study of new products is a cake that can be sliced a good many ways. You can study, for example, the way in which consumers switch from one brand to another. You might concentrate your attention on the size of the purchase of some group of consumers. You may prefer to study how long it takes a user to get around to buying again. You may be old-fashioned, or just artistic, and keep a sales chart over your desk. The authors prefer to study the expansion of new users and the time interval between repeat purchases.

The interesting features of any method of prediction in marketing or elsewhere are of three kinds. First there are some assumptions about the way the system, in this case the market, operates. If you keep up with the Joneses you will call this the mathematical model. Then you need some minimum amount of information to make any prediction at all. Thirdly, you and everyone else wants to know how well the prediction worked out.

In this article it is assumed that every new product is born with some upper limit to the number of its users. This upper limit does not stay the same but slowly and uniformly increases with time. It is also assumed that in any time interval, the number of extra new users gained will be proportional to the number of people represented by the upper limit of penetration who had not previously tried the product. These assumptions are sufficient to predict for any product the familiar growth curve looking like the top half of a martini glass on its side.

The information we need for prediction is then: an upper limit to the eventual penetration of the brand; a constant to indicate how this ceiling will increase with time; and the proportion of available non-users who will be persuaded to use the product in any time period. With all these we could predict the proportion of people using the brand at any time in the future, given some broad assumptions about the rest of the world staying much as it is and your competition lying down under your efforts.

Unfortunately the way in which this minimum information is put together is not quite clear. Obviously if the product has been around a little time and the rate at which it is gaining new customers is known, then the upper limit of its penetration is easily estimated. The authors' statement, "Experience with a large number of earlier new products is used to predetermine the general functional form or shape of the cumulative penetration as a function of time," seems to indicate that they, like the rest of us, draw some free-hand curves now and then.

If we accept these predictions of the number of users we can pass to the prediction of sales volume. For this the authors have studied the number of users who repeat their purchases once, twice, three times and so on in some definite time period. Their experience confirms what we might expect, that the average size of a repeat purchase is somewhat larger than the initial purchase. Looking ahead, they predict from the method previously described the number of new purchasers, and from their observations they predict how many of these have become repeat purchasers, buying the product various numbers of times. Together with the average purchase size this, of course, all adds and multiplies up to a prediction of sales volume.

Finally the hardest question: How good is the prediction? The illustration given in the article is a worked example showing a 23 per cent predicted loss compared to an actual loss of 27 per cent. This, the authors say, is not the pick of the crop. They might well have said that this is as much a prediction of the stability of our world as it is of the success of a new product.

This paper is a clear account in understandable terms of a method with strong intuitive appeal.

—GWYN COLLINS

Measuring the National Image

United States Information Agency—Office of Research and Analysis. *Post-Summit Trends in British and French Opinion of the United States and the Soviet Union, June 1960.* In the New York *Times*, October 27, 1960.

The U. S. Government, like many large corporations, carries out research to evaluate the effectiveness of its promotional program. The United States Information Agency, for example, regularly conducts surveys of foreign public opinion of the USA and its "competitor"—the Soviet Union. While these surveys are not normally published, two newspapers have recently printed their full texts.

The present survey, conducted shortly after the summit meetings, was not a full scale operation as was a pre-summit survey, but included only a limited group of selected questions as "riders" to ongoing research. The British study was based on 1,150 cases interviewed between May 21 and 24; the French study consisted of 1,000 cases interviewed between May 23 and 31. No further details on the sample plan are given, but interested readers are invited to write USIA for further details.

USIA notes that in the wake of the summit failure, both the USA and USSR suffered major losses in general standing. While the USSR is blamed most for the collapse of the summit conference, a majority blame both sides in part. Favorable opinion of the USA as a country and of President Eisenhower has declined according to the report, but remains far ahead of Soviet Russia and Premier Khrushchev, both of whom also suffered losses in esteem. Awareness of the sequence of three USA space successes was limited in Great Britain, and in France public opinion put the USSR overwhelmingly ahead in space development.

An interesting feature in reporting the results of these surveys is the use of a "net favorable" index: favorable answers minus unfavorable answers. According to USIA such an index often provides a useful basis of comparison, since it helps take into account the extent to which favorable responses are off-set by unfavorable responses. The meaningfulness of off-setting favorable by unfavorable responses, however, would require that the degree of total favorableness equal the degree of total unfavorableness. Even when degrees of favorable and unfavorable are available, i.e., "very good" and "good," or "very bad" and "bad," all good replies are apparently combined and subtracted from all bad replies; neutral replies, or no replies, are ignored. On this basis a shift in one country's standing from bad to neutral, or no replies, would show a gain in standing while a shift in the other country's standing from "good" to "very good" would show no gain. Even if "very good" were weighted twice as much as "good," we would have no guarantee that the distance from "good" to "very good" would be the same for any individual as his distance from "very bad" to "bad." It would seem that any algebraic manipulation of such an index could mislead the very person for whom it was designed, the busy executive not skilled in research techniques.

In most cases the questions asked of the respondent are reported along with the results. The procedure used in asking these, however, is not always given and it is impossible to tell whether precautions were taken against influencing the respondent by the order of listing the countries, the order of the alternatives, etc.

Standardized questions were asked with responses limited to several stated alternatives. Although a "no opinion" category was provided to prevent forcing the respondent to make a selection, this category has apparently received little thought in the analysis. In determining the reaction to the U-2 incident, the respondent was asked: "Have you heard or read about the recent incident in which an American plane, *said to have been spying*, was shot down by the Russians?" After being told that the plane *was spying* the respondent is then asked whether or not his opinion of the USA has been affected, whether the act was justified, etc. This one question, if asked early in the study, might have influenced the respondent's replies to all subsequent questions. In another case, the respondent was asked: "Is America doing all it should do to prevent a new world war?" The use of absolutes such as "all" generally encourages a negative response.

Public opinion is terribly difficult to analyze. Many factors influence attitudes, tools of measurement are complicated and at best one is never sure just what is being measured. To interpret such research done in foreign languages compounds the difficulties many times. USIA has nonetheless shown where something was amiss and provided pointers for remedial action.

—GEORGE M. SHIREY

PUBLICATIONS RECEIVED

MARY E. MCMAHON
ARF Librarian

Our collections have been enriched more rapidly than usual of late, partly due to the publication of this *Journal*, and partly because of the issuance of ARF's *Sources of Published Advertising Research*. Several book publishers and university presses have already supplied the *Journal* with review copies of recent books; many more have promised to do so as volumes of interest to advertising and marketing people become available.

The gathering of all materials listed in *Sources*—bibliographies, abstracting periodicals, directories, and lists of special libraries and professional bibliographers—began last December. Many, of course, were already on hand. Searches have been instituted for more elusive items and those now out of print. New publications of the types mentioned above will be acquired as they appear, and annotated in subsequent editions of *Sources*.

Thus ARF's library will probably be the most complete repository of secondary sources of advertising research. Most of these items are available on loan to ARF members, who are encouraged to use our library facilities.

Few primary sources such as books or articles are included in *Sources of Published Advertising Research*. Rather, it serves as a guide to the forms in which primary sources are collected and organized. Hopefully, *Sources* will remind researchers of existing reference tools and indicate new uses for them. It should also assist librarians in the accumulation of additional reference works.

Most of our new acquisitions are listed below. These are available from their respective publishers, or on loan from ARF.

Abelson, Herbert I., *Persuasion: How Opinions and Attitudes Are Changed*. New York: Springer Publishing Company, Inc., 1959. 118pp. $3.75.

*Audits and Surveys Company, Inc. and Erdos & Morgan Research Service, *The Third National Study of Sports Illustrated Subscriber Households*. New York: Time, Inc., 1960. 124pp. Free.

* Done in consultation with ARF.

Baylor, Aly, *A Motivational Research on the Greater Los Angeles Television Audience*. Los Angeles: Department of Economics and Marketing, Woodbury College, 1960. 10pp. Free.

Bristol, Lee H., Jr. (Editor), *Developing the Corporate Image: A Management Guide to Public Relations*. New York: Charles Scribner's Sons, 1960. 298pp. $5.95.

Burke Marketing Research, Inc., *A Chart Summary: The Effectiveness of Color vs. Black-and-White Television Advertising*. Cincinnati: Crosley Broadcasting Corporation, May 4, 1960. 11pp. Free.

Burke Marketing Research, Inc., *The Effectiveness of Color vs. Black-and-White Television Advertising*. Cincinnati: Crosley Broadcasting Corporation, May 4, 1960. 13pp. Free.

Cheskin, Louis, *Why People Buy: Motivation Research and Its Successful Application*. New York: Liveright Publishing Corporation, 1959. 348pp. $5.00.

Chicago Sun-Times, Chicago Daily News, *1960 Chicago Sun-Times, Chicago Daily News Consumer Analysis*. Chicago: Field Enterprises, Inc.—Newspaper Division, 1960. 230pp. Free.

Deming, W. Edwards, *Sample Design in Business Research*. New York: John Wiley & Sons, Inc., 1960. 517pp. $12.00.

Des Moines Register and Tribune, *Des Moines Register and Tribune Brand Inventory in the Homes of Iowa Subscribers*. Des Moines: Des Moines Register and Tribune, 1960. Over 100pp. Free.

Dichter, Ernest, *The Strategy of Desire*. Garden City, New York: Doubleday & Company, Inc., 1960. 314pp. $3.95.

The Dispatch Printing Company, *The 1960 Columbus Dispatch, Columbus Citizen-Journal Consumer Analysis*. Columbus: The Dispatch Printing Company, 1960. 110pp. Free.

Dix, Arthur H., *How to Measure the Value of Advertising*. New York: Conover-Mast Publications, Inc., 1960. 33pp. Free.

Dominick, Bennett A., Jr., *Research in Retail Merchandising of Farm Products—Appraisal of Methods and Annotated Bibliography*. Washington, D. C.: Agricultural Marketing Service, U. S. Department of Agriculture, July 1960. 46pp. Free.

Duluth Herald, Duluth News-Tribune, *The 1960 Consumer Analysis of the Duluth-Superior Market*. Duluth: Duluth Herald, Duluth News-Tribune, 1960. 98pp. Free.

*Erdos and Morgan, *The Esquire Subscriber's Leisure Activities*. New York: Esquire, August 1960. 48pp. Free.

Fawcett Publications, Inc. Research Department, *Men's Apparel Survey*. New York: True, The Man's Magazine, 1960. 47pp. Free.

Fenton, John M., *In Your Opinion . . . The Managing Editor of the Gallup Poll Looks at Polls, Politics, and the People from 1945 to 1960*. Boston: Little, Brown and Company, 1960. 220pp. $3.95.

Good Housekeeping, *Dessert Survey: A Good Housekeeping Consumer Panel Report*. New York: The Hearst Corporation, 1960. 86pp. Free.

Gunther, Edgar and Frederick A. Goldstein, (Editors). *Current Sources of Marketing Information: A Bibliography of Primary Marketing Data*. Chicago: American Marketing Association, August 1960. 119pp. $2.00 to AMA members; $4.00 to non-members.

Hanan, Mack, *The Pacifiers: The Six Symbols We Live By*. Boston: Little, Brown and Company, 1960. 306pp. $5.00.

Hattwick, Melvin S., *The New Psychology of Selling*. New York: McGraw-Hill Book Company, Inc., 1960. 276pp. $5.95.

Havas, Nick, Violet Davis Grubbs and Hugh M. Smith, *Combining Kinds of Retailer Promotions: Effect on Sales of Selected Food Products*. Washington, D. C.: Agricultural Marketing Service, U. S. Department of Agriculture, August 1960. 16pp. Free.

Havas, Nick, Peter L. Henderson, Chester S. Parsons and Paul Schaffer, *Displaying Fruit in Various Types of Packages and in Bulk . . . Costs and Effects on Quality and Sales*. Washington, D. C.: Agricultural Marketing Service, U. S. Department of Agriculture, August 1960. 22pp. Free.

Havas, Nick and Hugh M. Smith, *Customers' Shopping Patterns in Retail Food Stores: An Exploratory Study*. Washington, D. C.: Agricultural Marketing Service, U. S. Department of Agriculture, August 1960. 14pp. Free.

Honolulu Star-Bulletin, *1960 Consumer Analysis of the Oahu Market (including Honolulu)*. Honolulu, Hawaii: Honolulu Star-Bulletin, 1960. 123pp. Free.

Howard, Ronald A., *Dynamic Programming and Markov Processes*. New York: John Wiley & Sons, Inc., 1960. 136pp. $5.75.

Kirkpatrick, C. A., *Advertising: Mass Communication in Marketing*. Boston: Houghton Mifflin Company, 1959. 638pp. $7.50.

Lockley, Lawrence C., *Use of Motivation Research in Marketing*. New York: National Industrial Conference Board, Inc., 1960. 27pp. $1.50 to NICB associates; $7.50 to non-associates.

Lohse, Charles F., *Creative Selling*. New York: Charles Scribner's Sons, 1960. 212pp. $4.95.

The Long Beach Independent, Press-Telegram, *1960 Consumer Analysis of the Long Beach, California Market*. Long Beach: The Long Beach Independent, Press-Telegram, 1960. 91pp. Free.

Marplan, *A Study of Farm Publications in Eight Midwestern States*. Chicago: Midwest Farm Paper Unit, 1960. 70pp. Free.

Mason, Edward S., (Editor), *The Corporation in Modern Society*. Cambridge: Harvard University Press, 1960. 335pp. $6.75.

McCarthy, E. Jerome, *Basic Marketing: A Managerial Approach*. Homewood, Illinois: Richard D. Irwin, Inc., 1960. 770pp. $10.60.

McClatchy Newspapers, *1960 Consumer Analysis: The Sacramento Bee, The Fresno Bee, The Modesto Bee*. Sacramento, California: McClatchy Newspapers, 1960. 264pp. Free.

McGraw-Hill Research, *The Purchasing Week Survey of Buying Influence Patterns*. New York: McGraw-Hill Publishing Co., Inc., 1960. 62pp. Free.

Meloan, Taylor W. and John M. Rathmell, *Selling: Its Broader Dimensions*. New York: The Macmillan Company, 1960. 409pp. $7.50.

National Industrial Conference Board, *A Graphic Guide to Consumer Markets*. New York: National Industrial Conference Board, 1960. 112pp. $3.50 to NICB associates; $35.00 to non-associates.

Nation's Business Research Department, *A Study of the Corporate Personality: How Business Executives Appraise Leading Corporations Holding Military Prime Contracts*. Washington, D. C.: Nation's Business, 1960. 107pp. Free.

Newark News, *The Newark News Consumer Analysis*. Newark, New Jersey: Newark News, 1960. 62pp. Free.

The Omaha World-Herald, *1960 Omaha World-Herald Consumer Analysis*. Omaha, Nebraska: The Omaha World-Herald, 1960. 112pp. Free.

Oregon Journal, *1960 Consumer Analysis of the Portland, Oregon Market*. Portland: Oregon Journal, 1960. 98pp. Free.

Organisation for European Economic Co-operation, *Market Research on a European Scale*. Paris: The European Productivity Agency of the Organisation for European Economic Co-operation, June 1960. 136pp. $2.00.

Politz, Alfred, Media Studies, *The Rochester Study*. Philadelphia: The Saturday Evening Post, August 1960. 83pp. Free.

The Providence Journal-Bulletin, *1960 Providence Consumer Analysis*. Providence, Rhode Island: The Providence Journal-Bulletin, 1960. 115pp. Free.

Rapoport, Anatol, *Fights, Games, and Debates*. Ann Arbor: The University of Michigan Press, 1960. 400pp. $6.95.

Richmond Times-Dispatch, The Richmond News Leader, *Households Reached by Radio and Television in Metropolitan Richmond*. Richmond, Virginia: Richmond Newspapers, Inc., June 1960. 7pp. Free.

The St. Paul Dispatch-Pioneer Press, *St. Paul Dispatch-Pioneer Press 1960 Consumer Analysis of the St. Paul ABC City Zone*. St. Paul, Minnesota: The St. Paul Dispatch-Pioneer Press, 1960. 214pp. Free.

San Jose Mercury and News, *1960 Consumer Analysis of the San Jose ABC City Zone*. San Jose, California: San Jose Mercury and News, 1960. 96pp. Free.

Sandage, C. H. and Vernon Fryburger (Editors), *The Role of Advertising: A Book of Readings*. Homewood, Illinois: Richard D. Irwin, Inc., 1960. 499pp. $8.00.

The Seattle Times, *The Seattle Times 1960 Consumer Analysis of Greater Seattle*. Seattle: The Seattle Times, 1960. 119pp. Free.

Simon, Herbert A., *The New Science of Management Decision*. New York: Harper & Brothers, 1960. 50pp. $2.50.

Smith, Hugh M., *Customer Traffic Research*. Washington, D. C.: Agricultural Marketing Service, U. S. Department of Agriculture, July 1960. 11pp. Free.

Social Research, Inc., *The Invisible Wall . . . The Social Factor in Mass Selling*. New York: True Story Women's Group, 1960. 22pp. Free.

Starch, Daniel and Staff, *A Study Among Farmers Actively Engaged in Farming in New York State*. Ithaca: American Agriculturist, 1960. 15pp. Free.

Television Advertising Representatives, Inc., *Viewing by Mothers of Young Children*. New York: Television Advertising Representatives, Inc., 1960. 19pp. Free.

Wallach, M. A., Research, Inc., *Summer Servings in Urban American Homes*. New York: American Can Company, undated. 31pp. Free.

Weiss, E. B., *Planning Merchandising Strategy for 1961-1965*. New York: Doyle Dane Bernbach, Inc., 1960. 73pp. Free.

The Wichita Eagle, *Wichita Eagle Consumer Analysis 1960*. Wichita, Kansas: The Wichita Eagle, 1960. 98pp. Free.

Yamanaka, Jiro, *A Method of Prediction of Readership Score (Newspaper Ads)*. Tokyo: Dentsu Advertising Ltd., 1960. 44pp. Free.

Uniquack on Advertising Research

Readers of the New York *Times* editorial page will recognize Uniquack as the computer that knows all the answers. With a low bow to its usual interrogator, James Reston, we publish herewith the transcript of a conversation between Uniquack and a representative of the advertising industry.

AD MAN: We've all heard of your talents, Uniquack, and appreciate this opportunity to use them on problems of advertising research. But I hope you won't mind if I begin with what may be an embarrassing question. How is it your fellow computers were so erratic in predicting the vote election night?

UNIQUACK: I feel no responsibility to defend my more fallible colleagues. They, poor fellows, are only as accurate as the programs and data they get fed.

A: But why did the odds fluctuate so wildly?

U: The trouble with you humans is that you keep confusing probabilities with amounts. Those were the odds that one man would win, even if by only one vote. Guesses about the margin of victory weren't tied to those odds.

A: I'm confused already. But we're getting off the track. My real question is: can the sales or profit resulting from advertising really be measured?

U: Certainly. As a matter of fact, it already has.

A: You don't say. By whom?

U: Don't you think that back in the days when the catalog was their only promotional medium, the big mail-order houses knew pretty well how much they earned per dollar spent on advertising space?

A: Maybe so, but you've taken an extreme case. There was only one advertising medium and only one place the consumer could buy. How about measuring the profitability of advertising a well-distributed product in mass media?

U: Same problem. Just more record keeping and busy work. And what are business machines for?

A: That's a laugh, coming from you. You can't mean that in these days of the integrated marketing plan, we can measure the effects of any part of the plan just by feeding accurate records into a machine! What about pricing, distribution, coupons, free samples, dealer promotions? What about the sales force?

U: What about it?

A: Well, it seems to me that if we're going to find out what sales result from advertising, we would have to stop all the other marketing activities that lead to sales. No firm is going to stop using most of its weapons just to find out how well *one* of them works. Or are you suggesting we all go back to the mail-order business and use only catalogs?

U: Not at all. You needn't control all the sales-causing factors as long as you measure how your total effort is divided among them—in each locality and in each time period. To make matters easier, you might record all the efforts in some common unit.

A: Such as?

U: As Alexander Woollcott replied to the woman's club president who said she didn't know how to thank him, "ever since the Phoeniceans invented money, that has been no problem."

A: Very funny. So we record dollars spent on each marketing action, in each place, during each period, and look for corresponding sales changes. Suppose we find that wherever and whenever advertising expenditures were increased, sales went up. How do we know that these sales weren't caused by competition relaxing, or the weather, or any other factor we can't control? Or even that we increased the ad budget because the sales increased?

U: You don't. The whole point of this record keeping is to permit you to design and conduct useful experiments.

A: Experiments? Have you seen our last earnings statement?

U: The smaller your profits, the greater may be your need to get the benefits of experimentation. Call it "planned variation" if you like. All it means is that you vary advertising in a few test markets chosen at random.

A: At random! That's not the way *we* plan variations in advertising between markets.

U: Then you'll never know the degree to which advertising caused whatever sales changes you later observe. When you pick randomly when and where you'll change your advertising expenditure, then the effects of all those "other factors" can't possibly bias your results except by chance. If you don't mind my saying so, it's also a way of eliminating "human" error.

A: I'll let that pass. But I'm beginning to see what you mean. If I leave it to the sales manager to pick the areas that get doubled advertising budgets, he might pick those where sales have been bad for other reasons. Then those other reasons would be all mixed up with the doubled advertising as causes of whatever happened.

U: Exactly.

A: But suppose just by accident we pick test markets that are affected by some other factor I not only can't control but also can't even measure?

U: You'll never know for sure. But you will know the odds of that happening by chance alone. If your experiment has lots of test markets, you can get those odds up pretty high.

A: Like those on election night?

U: Yes, but remember that you can test the results of *your* predictions more often than just every four years.

A: That's some comfort, anyway. If it's as easy as you make out, why hasn't everyone tried experiments with advertising expenditures?

U: Now you're getting out of my field. You fellows set the goals, we just do the adding and subtracting. Some days I wish I knew just what your goals are.

A: Hmm. Well, thanks for a most instructive interview, Uniquack. Sorry to cut it short, but it's lunch time and I'm hungry. But then you wouldn't know what that means.

—CHARLES K. RAMOND